LEÇONS

SUR LA

THÉORIE DES FONCTIONS

DU MÊME AUTEUR

Librairie FÉLIX ALCAN

Le Hasard, 1914 (7e édition, 1923).
L'Espace et le Temps, 1922 (7e édition, 1925).
L'Aviation (en collaboration avec MM. Paul PAINLEVÉ et Ch. MAURAIN, 8e édition, 1925).

Librairie ARMAND COLIN

Cours élémentaire de Mathématiques (*Arithmétique*, *Algèbre*, *Géométrie*, *Trigonométrie*).
Probabilités. Erreurs (en collaboration avec M. Robert DELTHEIL).

Librairie HERMANN

Éléments de la Théorie des Probabilités (3e édition, 1924).

Librairie GAUTHIER-VILLARS

Leçons sur les Fonctions entières, 1900 (2e édition, 1920).
Leçons sur les Séries divergentes (2e édition, 1927).
Leçons sur les Séries à termes positifs, 1902.
Leçons sur les Fonctions méromorphes, 1903.
Leçons sur les Fonctions de variables réelles (2e édition, 1928).
Leçons sur la Théorie de la croissance, 1909.
Introduction géométrique à quelques Théories physiques, 1914.
Leçons sur les Fonctions monogènes uniformes d'une variable complexe, 1917.
Problèmes et Méthodes de Théorie des fonctions, 1922.
Traité du Calcul des Probabilités et de ses applications (en cours de publication).

Librairie VUIBERT

Introduction à l'étude de la Théorie des nombres et de l'Algèbre supérieure (en collaboration avec M. Jules DRACH), 1895 (épuisé).

Librairie ALBIN MICHEL

Principes d'Algèbre et d'Analyse, 1925.

COLLECTION DE MONOGRAPHIES SUR LA THÉORIE DES FONCTIONS
PUBLIÉE SOUS LA DIRECTION DE M. ÉMILE BOREL

LEÇONS
SUR LA
THÉORIE DES FONCTIONS

(*Éléments et principes de la théorie des ensembles : applications à la théorie des fonctions*)

PAR

ÉMILE BOREL

PROFESSEUR A LA FACULTÉ DES SCIENCES DE PARIS
MEMBRE DE L'INSTITUT

TROISIÈME ÉDITION

PARIS

GAUTHIER-VILLARS ET Cie, ÉDITEURS

LIBRAIRES DU BUREAU DES LONGITUDES, DE L'ÉCOLE POLYTECHNIQUE

55, Quai des Grands-Augustins.

1928

PRÉFACE DE LA TROISIÈME ÉDITION

La deuxième édition de ce Livre a été, malgré les années de guerre, épuisée plus rapidement que la première; les raisons qui m'ont décidé à ne pas remanier le texte primitif et que j'ai indiquées dans la Préface de la deuxième édition, subsistent intégralement [1]. J'ai seulement ajouté une septième Note, consacrée aux discussions récentes sur la logique empirique. Je tiens à remercier MM. Robert Wavre, Paul Lévy, Alfred Errera et Barzin de m'avoir autorisé à reproduire leurs articles dans cette Note.

5 août 1927.

[1] Au sujet des questions traitées dans la Note V (*Probabilités dénombrables*), je prends la liberté de renvoyer à mon *Traité du Calcul des Probabilités et de ses applications* (Gauthier-Villars) et en particulier au fascicule I du Tome II (*Application à l'arithmétique et à la théorie des fonctions*).

PRÉFACE DE LA DEUXIÈME ÉDITION

La première édition de ce Livre a été rédigée d'après des conférences faites à l'École Normale pendant l'année scolaire 1896-1897. En me confiant cet enseignement, Jules Tannery m'a donné une marque de confiance dont j'ai immédiatement senti tout le prix; l'intérêt pris à ces leçons par leurs auditeurs m'a encouragé à rédiger ce petit livre; ainsi est née cette Collection de Monographies sur la théorie des fonctions qui s'honore maintenant de tant d'éminents collaborateurs. C'est donc à Jules Tannery que cette collection doit d'exister; c'est à lui tout d'abord que doit aller la reconnaissance de ceux qui pensent qu'elle n'a pas été inutile.

Je me suis demandé s'il convenait de remanier cette nouvelle édition; mais je me suis rapidement aperçu que je serais conduit à reprendre des matières traitées dans d'autres Ouvrages de la Collection; il m'a donc paru préférable de conserver à ce Livre le caractère élémentaire qui en fait une sorte d'introduction à plusieurs de ces Ouvrages (1); le texte des six Chapitres et des trois Notes a donc été reproduit sans modification (2). J'ai, par contre, ajouté trois nouvelles Notes qui doublent presque les dimensions du livre et au sujet desquelles il me paraît utile de donner quelques explications.

---

(1) Je n'ai pas cru devoir ajouter une liste des travaux qui se rattachent aux questions traitées ici et qui ont été publiés depuis la première édition; voir : *Recherches contemporaines sur la théorie des fonctions* dans l'*Encyclopédie des Sciences mathématiques*.

(2) Un seul changement a été fait; pour la définition des ensembles parfaits et fermés, où le langage proposé par M. Jordan avait introduit quelque flottement, j'ai adopté la terminologie qui paraît avoir actuellement prévalu sans contestation.

Dans la Note IV, j'ai réuni divers articles relatifs aux discussions sur les principes de la théorie des ensembles ; je remercie MM. Baire, Hadamard et Lebesgue de m'avoir autorisé à reproduire la correspondance que nous avons échangée en 1904 et qui a été publiée dans le *Bulletin de la Société mathématique*. La comparaison entre la Note IV et les pages qui la précèdent montre que ma pensée a évolué sur plusieurs points ; il m'a semblé que je ne devais pas dissimuler cette évolution, d'autant que j'ai parfois constaté chez ceux qui étudient pour la première fois la théorie des ensembles une évolution analogue dans l'interprétation des principes ; il y a là des questions assez complexes pour qu'il ne soit pas inutile d'en mettre successivement en évidence divers aspects un peu différents. Ces discussions sont d'ailleurs peut-être plus intéressantes pour les philosophes que pour les mathématiciens ; j'espère que les philosophes qui s'intéressent aux principes et qui veulent bien citer parfois ce Livre, me sauront gré d'avoir réuni des articles dispersés dans de nombreux Recueils et d'y avoir joint une exposition inédite de la « démonstration » de M. Zermelo qui a joué un rôle dans beaucoup de ces controverses.

La Note V est consacrée aux probabilités dénombrables, qui se rattachent étroitement à plusieurs des questions étudiées précédemment, notamment à la nature arithmétique des incommensurables et à la théorie de la mesure.

Dans la Note VI [1] enfin, j'ai essayé de développer la théorie de la mesure et de l'intégration en restant au point de vue que j'avais adopté pour définir les ensembles mesurables. On sait quelle extension M. Lebesgue a donnée à ma définition et quelle riche moisson de résultats il en a tirés ; quelques-uns de ces résultats ont été exposés par lui-même dans deux livres de cette Collection : *Leçons sur l'intégrale définie* et *Leçons sur les séries trigonométriques*. Il m'a semblé qu'il n'était peut-être pas inutile de faire voir comment les définitions constructives dont j'étais parti conduisent par une voie très élémentaire et très simple à des résultats pratiquement aussi généraux, tant pour

[1] Cette Note reproduit, avec quelques modifications, un Mémoire que j'ai publié en 1912 dans le *Journal* de M. Jordan.

l'intégration que pour la mesure. Je serais heureux si cette Note VI pouvait être, pour beaucoup de lecteurs, une introduction à l'étude des belles recherches de M. Lebesgue, dont l'importance capitale est chaque jour mieux comprise et dont les conséquences sont loin d'être épuisées.

Je tiens enfin à renouveler à M. Gauthier-Villars l'expression de ma gratitude pour la part qu'il a eue dans la création et le développement de cette Collection de monographies et en particulier pour le désintéressement avec lequel il m'a permis de lui maintenir le niveau scientifique élevé que doivent avoir des livres dont le but principal est de provoquer de nouvelles recherches.

15 juillet 1914.

PRÉFACE DE LA PREMIÈRE ÉDITION

Le titre que j'ai cru devoir adopter est assez vague pour qu'il ne me semble pas inutile de donner quelques explications sur l'objet de ces *Leçons*. Les dimensions de ce petit livre me dispensent, je l'espère, de dire que ce n'est point un Traité complet sur une branche des Mathématiques dont l'étendue s'accroît chaque jour. Ce n'est pas non plus un exposé nouveau des principes de la théorie; ces principes, rendus classiques en France par la publication du célèbre Cours autographié de M. Hermite, ont été exposés depuis dans plusieurs Ouvrages, qu'il n'y a pas lieu de remplacer.

Mais la lecture des Mémoires originaux devient chaque jour plus difficile pour celui qui connaît seulement de la Théorie des fonctions les parties regardées actuellement comme classiques; il m'a dès lors semblé qu'on pouvait chercher à faire œuvre utile en tentant d'exposer, d'une manière élémentaire, certaines recherches qui, bien que relativement récentes, prennent chaque jour une importance plus considérable. De ce nombre est la Théorie des ensembles; c'est à elle qu'est consacré cet Ouvrage. J'ai tenu cependant à lui donner le titre de *Leçons sur la Théorie des fonctions*, car, en parlant des ensembles, j'ai cherché à ne jamais perdre de vue les applications.

Ce n'est pas que je méconnaisse le très haut intérêt que présente par elle-même la Théorie des ensembles; mais il m'a paru qu'il y avait lieu de distinguer nettement cet intérêt philosophique de l'utilité pratique de la théorie, c'est-à-dire de son lien avec d'autres parties des Mathématiques. Aussi ai-je laissé de côté bien des résultats intéressants obtenus par divers géomètres au sujet des ensembles, parce que je n'aurais pas pu en donner d'applications ici même.

Les trois premiers Chapitres sont un exposé des éléments de la Théorie des ensembles; j'ai cherché à les rendre aussi accessibles que possible, en supposant chez le lecteur un minimum de connaissances.

Les trois derniers Chapitres contiennent des applications à la Théorie des fonctions; j'ai, cette fois, supposés connus certains résultats qui sont établis dans l'un quelconque des Ouvrages cités il y a un instant et semblent, par suite, pouvoir être regardés comme classiqnes.

Je n'ai d'ailleurs pas cherché à remplacer la lecture des Mémoires originaux, mais seulement à la faciliter; aussi ai-je laissé des lacunes qu'il aurait été aisé de combler en transcrivant presque textuellement un certain nombre de pages de tel ou tel Mémoire; il y a toujours avantage, pour le lecteur qui désire approfondir une question, à recourir lui-même au Mémoire original.

Des Notes sont consacrées à quelques questions qui se rapportent aux principes de la Théorie des ensembles et dont la discussion aurait alourdi inutilement les premiers Chapitres.

15 février 1898.

INDEX

LEÇONS

SUR LA

THÉORIE DES FONCTIONS.

CHAPITRE I.

NOTIONS GÉNÉRALES SUR LES ENSEMBLES.

I. — La notion d'ensemble.

Nous ne chercherons pas à donner une définition du mot « ensemble » : il nous paraît qu'il y a là une notion suffisamment primitive pour qu'une *définition* en soit au moins inutile ; on peut seulement se proposer d'éclaircir par des exemples le sens de ce mot, mais ces exemples ne donneraient pas l'*idée* d'ensemble à celui qui ne l'aurait pas. On peut dire que l'on entend par *ensemble* une collection d'un nombre fini ou infini d'objets d'ailleurs quelconques ; ce mot *collection* doit donc être pris dans le sens le plus large possible ; par exemple, on dira l'*ensemble des nombres rationnels*, l'*ensemble des nombres positifs*, l'*ensemble des fonctions continues d'une variable réelle*, l'*ensemble des points situés sur un cercle donne et dont le sinus est incommensurable*, etc. Nous supposons donc que nous avons l'*idée* d'ensemble; mais cette idée n'a point encore une netteté suffisante pour pouvoir figurer dans un raisonnement mathématique; le but des quelques pages qui vont suivre est précisément de chercher à donner à la notion d'ensemble la précision qui est nécessaire pour qu'on puisse l'utiliser dans des recherches rigoureuses.

La première difficulté qui se présente est la suivante : lorsqu'un

ensemble est composé d'une infinité d'éléments, quand doit-on le considérer comme donné [1]? Il est bien clair, en effet, que, si le nombre des éléments est *fini*, donner un ensemble, c'est donner tous ses éléments, sans en excepter un seul. Peut-on adopter une telle définition lorsque le nombre des éléments est infini : l'expression *donner une infinité d'éléments* doit-elle être regardée comme ayant un sens?

II. — La notion de puissance.

On rencontre cette question pour la première fois dans les éléments, lorsqu'on aborde l'étude des séries [2] et l'on adopte d'habitude la définition suivante : une série infinie $u_1, u_2, \ldots, u_n, \ldots$ est *donnée*, si l'on sait calculer le *terme général* u_n lorsqu'on connaît son rang n. En d'autres termes, u_n est une fonction donnée de n. D'ailleurs, u_n peut être donné explicitement en fonction de n, ou bien être seulement déterminé par une loi quelconque,

---

(1) M. G. Cantor, dans un des Mémoires de haute importance qu'il a publiés sur la théorie des ensembles, donne la définition suivante : « Je dis qu'un ensemble d'éléments appartenant à une sphère abstraite quelconque est *bien défini* quand, par suite du principe logique du troisième exclu, on peut le considérer déterminé de telle facon que : 1° un objet quelconque appartenant à cette sphère abstraite étant choisi, l'on puisse regarder comme *intrinsèquement* déterminé s'ils sont égaux ou non, malgré les différences qui peuvent se présenter dans la manière dont ils sont donnés.

» En fait, on ne ponrra pas généralement effectuer d'une manière sûre et précise les déterminations en question *avec les méthodes* ou *les moyens dont on dispose;* mais là n'est pas la question.... Pour éclaircir ceci, je rappelle la définition du système de tous les nombres algébriques; on peut, sans aucun doute, concevoir être déterminé intrinsèquement si un nombre η, choisi à volonté, appartient ou non aux nombres algébriques; néanmoins, le problème qui consiste à trouver cette détermination par rapport à un nombre donné η est souvent, comme on le sait, un des plus difficiles. » (*Acta mathematica*, t. II, p. 363.) On verra que nous nous plaçons à un point de vue très différent de celui de M. G. Cantor, dont nous avons tenu à citer textuellement ce passage, car il nous a paru très intéressant, mais en même temps assez difficile à comprendre pour que nous ayons craint d'en altérer le sens en cherchant à le résumer ou à l'interpréter.

(2) On la rencontrerait aussi dans la définition des nombres incommensurables, par une fraction décimale ou par tout autre moyen; mais on a alors moins l'habitude de la mettre en évidence; nous reviendrons plus loin sur ce point.

par exemple par une relation de récurrence au moyen des termes précédents. Dans tous les cas, on connaît un moyen de calculer successivement, les uns après les autres, tous les termes de la série, en les obtenant d'ailleurs chacun une seule fois [1].

Nous sommes ainsi amenés à proposer la définition suivante : *Nous dirons qu'un ensemble est donné lorsque, par un moyen quelconque, on sait en déterminer tous les éléments les uns après les autres, sans en excepter un seul et sans répéter aucun d'eux plusieurs fois.* Cette définition, à laquelle nous avons été naturellement conduits, paraîtra sans doute claire au lecteur non encore familier avec cette théorie; mais l'une des idées que nous serions le plus heureux de faire acquérir à celui qui veut arriver à penser par lui-même sur la théorie des fonctions, c'est que, dans toutes les questions où intervient l'infini, il faut se méfier extrêmement de la prétendue clarté : rien n'est plus dangereux que de se payer de mots en pareille matière. C'est ce que nous allons constater en analysant cette définition et nous serons ainsi conduits à plusieurs notions importantes.

Que veulent dire, en effet, ces mots : *Déterminer tous les éléments les uns après les autres?* Reprenons l'exemple de la série u_n; donner, les uns après les autres, les éléments de l'ensemble u_n, a ici un sens parfaitement clair; c'est donner d'abord u_1, puis u_2, puis u_3, puis u_4, ... Mais, si nous prenons un autre exemple, et si nous considérons, pour fixer les idées, l'ensemble de tous les points d'une circonférence, que sera-ce que *les donner tous, les uns après les autres?* Dans quel ordre faut-il les prendre pour être assuré de les avoir tous? Sans étudier encore d'une manière approfondie cette embarrassante question, sur laquelle nous reviendrons dans les Notes, nous pourrions être tentés de répondre : Pour avoir tous les points d'un cercle, il

---

[1] On pourrait aussi imaginer que u_n est une fonction de n que l'on conçoit, sans qu'elle soit explicite; au sujet d'une telle dépendance (entre u_n et n) qui ne correspond à aucune *loi* mathématique, on consultera avec le plus grand fruit un intéressant article de M. Jules Tannery dans la *Revue générale des Sciences* (28 février 1897). La définition donnée par M. G. Cantor dans la note de la page précédente se rattache au même ordre d'idées. Mais nous n'insistons pas sur ce point, car, sans en méconnaître le très haut intérêt, nous devons constater qu'il a, actuellement, peu d'importance au point de vue purement mathématique.

suffit de le tracer avec un compas : la pointe du compas passera successivement par tous ces points, sans en excepter un seul. Il n'y a pas d'inconvénient à accepter provisoirement cette réponse, mais nous ne pouvons manquer d'être frappés des différences essentielles qu'elle présente avec celle que nous avons pu faire dans le cas de la série des u_n. Pour ne signaler actuellement qu'une de ces différences : dans la série un élément est immédiatement précédé (et suivi) d'un autre élément parfaitement déterminé ; il n'en est nullement de même sur le cercle : un point n'a pas de point consécutif.

Qu'est-ce qui facilite notre conception de l'*ensemble* formé par les termes de la série? Il est aisé de se rendre compte que c'est le fait d'avoir pu donner un *indice* à chacun de ces termes, de sorte que chacun d'eux corresponde à un nombre entier positif déterminé n. Dès lors, la conception que nous avons de l'*ensemble* des u_n est aussi claire que notre conception de l'*ensemble des nombres entiers positifs* $1, 2, 3, \ldots, n, \ldots$. Or, ce dernier ensemble nous est très familier (1); nous avons certainement le droit de le regarder comme connu. C'est grâce à cette connaissance que les mots : *l'ensemble des u_n est donné* ont pour nous un sens bien précis.

Si nous cherchons à analyser le procédé par lequel nous nous sommes donné cet ensemble des u_n, voici ce que nous constatons : nous sommes partis d'un ensemble que nous avons considéré comme *donné :* l'ensemble des nombres entiers positifs; puis, à chaque élément n de cet ensemble *donné*, nous avons fait correspondre (2) un élément de l'ensemble qu'il s'agissait de définir et qui s'est trouvé ainsi *donné* d'une manière précise. Nos deux ensembles, l'ensemble des nombres entiers positifs, que nous désignerons par E, et l'ensemble des u_n, que nous désignerons par U, sont donc liés d'une manière très simple. A chaque élé-

---

(1) Ce n'est point ici le lieu d'étudier les raisons pour lesquelles cet ensemble nous apparaît comme bien connu, ni les procédés par lesquels notre esprit a acquis cette connaissance. On trouvera un résumé et une discussion des diverses opinions émises à ce sujet, ainsi que des renseignements bibliographiques, dans l'intéressante thèse de M. Couturat : *De l'Infini mathématique.* Voir aussi l'article de M. Jules Tannery, déjà cité.

(2) *Voir* la note de la page 3.

ment (n) de E correspond un élément (u_n) de U et un seul, et réciproquement à chaque élément (u_n) de U correspond le seul [1] élément (n) de E.

Il résulte d'ailleurs de là qu'à des éléments différents de l'un des ensembles correspondent des éléments différents de l'autre. On peut exprimer cette relation entre E et U en disant qu'il existe entre leurs éléments une correspondance *univoque* et *réciproque;* nous conviendrons de dire, plus brièvement, que les deux ensembles ont même *puissance*. La notion de puissance, introduite par M. G. Cantor, est fondamentale dans la théorie des ensembles; comme nous venons de le voir, *deux ensembles sont dits avoir même puissance lorsqu'on peut établir entre leurs éléments une correspondance telle, qu'à tout élément de chacun d'eux corresponde un élément et un seul de l'autre.*

Il est clair que, si deux ensembles A et B ont chacun même puissance qu'un troisième ensemble C, A a même puissance que B. Au lieu de dire que deux ensembles ont même puissance, on dit aussi que leurs puissances sont égales, de sorte que nous aurions pu exprimer le fait précédent en disant que deux puissances égales à une même troisième sont égales entre elles; mais nous préférons éviter ce langage, qui a l'inconvénient de suggérer une analogie entre les puissances et les grandeurs mesurables, analogie qui n'est pas suffisamment justifiée; nous reviendrons d'ailleurs sur ce point (Note I).

En définitive, pour nous donner l'ensemble U, nous sommes partis d'un ensemble E, de même puissance que U, supposé donné. Il est clair que, d'une manière générale, si l'on regarde un ensemble quelconque comme donné, on pourra regarder comme donnés tous les ensembles de même puissance qu'on peut en déduire par un procédé analogue à celui qui, appliqué à E, nous a donné U. Par exemple, regardons comme donné l'ensemble formé de tous les nombres incommensurables compris entre o et 1; soit α l'un quelconque d'entre eux; prenons d'autre part, sur une circonférence de rayon R, une origine des arcs A et marquons tous les points M tels que l'arc AM soit égal à $2\alpha\pi R$. Il est clair que

---

[1] Bien entendu, on suppose que les u_n sont tous différents, ou du moins pensés comme tels.

l'ensemble des points M sera ainsi défini : c'est l'ensemble des extrémités des arcs incommensurables avec la circonférence.

Si donc on se propose d'étudier les ensembles dans leur généralité, la première question qui se pose est de savoir combien il y a de puissances différentes et la seconde, de connaître un ensemble possédant chacune des puissances existantes. Mais, avant d'aborder ces difficiles questions, et aussi celle de savoir si l'on peut se donner des ensembles par un procédé différent de celui que nous avons indiqué, il ne sera pas inutile d'étudier les ensembles qui ont même puissance que l'*ensemble* E *des nombres entiers positifs*. Nous reconnaîtrons, en effet, que cette puissance est la plus simple possible (¹) et que les ensembles qui la possèdent ont des propriétés caractéristiques et importantes.

III. — Les ensembles dénombrables.

Ces ensembles sont dits *ensembles dénombrables;* voici ce que signifie cette expression : si un ensemble a même puissance que l'ensemble E, nous pouvons toujours convenir de désigner par u_n l'élément qui correspond au nombre entier n; nous pourrons alors ranger les éléments de l'ensemble dans une suite

$$u_1,\quad u_2,\quad u_3,\quad \ldots,\quad u_n,\quad \ldots,$$

tout à fait analogue à celle des nombres entiers, dans laquelle chacun des éléments a un rang déterminé, un élément qui le précède et un élément qui le suit immédiatement.

Nous allons indiquer d'abord des classes importantes d'ensembles dénombrables; nous montrerons ensuite qu'il y a des ensembles qui ne sont pas dénombrables, c'est-à-dire qui n'ont pas même puissance que E.

Par définition même, est dénombrable tout ensemble dont chaque élément est affecté d'un indice, les indices prenant toutes les valeurs entières et positives. Supposons maintenant que nous

---

(¹) Nous laissons de côté les ensembles composés d'un nombre fini d'éléments pour eux la notion de puissance se confond avec la notion de nombre : ils ont même puissance lorsque leurs éléments sont en même nombre.

ayons des éléments u_α, tels que l'indice α prenne seulement des valeurs entières positives, mais sans les prendre toutes; d'ailleurs, bien entendu, deux éléments différents ont toujours des indices différents. Dans ces conditions, l'ensemble des u_α est encore dénombrable; en effet, les indices α étant des nombres entiers positifs, nous pouvons les ranger par ordre de grandeur croissante, désigner le plus petit par α_1, celui qui est le plus petit parmi les autres par α_2, le suivant par α_3, etc. Si nous posons ensuite

$$u_{\alpha_1} = v_1, \qquad u_{\alpha_2} = v_2, \qquad u_{\alpha_3} = v_3, \qquad \ldots,$$

nous reconnaîtrons que notre ensemble, pouvant se mettre sous la forme

$$v_1, \quad v_2, \quad v_3, \quad \ldots,$$

est *dénombrable*. Par exemple, on pourrait supposer que les indices α sont tous les nombres pairs, ou bien tous les nombres premiers, etc. On reconnaît ainsi que l'ensemble des nombres pairs, par exemple, a *même puissance* que l'ensemble E de tous les nombres entiers, et, effectivement, si l'on fait correspondre à chaque nombre pair le nombre dont il est le double, on aura réalisé une correspondance univoque et réciproque entre ces deux ensembles. On voit combien la notion de puissance se distingue de la notion de nombre; l'axiome arithmétique : le tout est plus grand que la partie, n'a pas d'analogue dans la théorie des puissances : l'ensemble des nombres pairs est une *partie* de E et cependant a même puissance que E [1].

Considérons maintenant un nombre limité d'ensembles dénombrables, trois pour fixer les idées :

$$\begin{array}{llll} u_1, & u_2, & u_3, & \ldots, \\ v_1, & v_2, & v_3, & \ldots, \\ w_1, & w_2, & w_3, & \ldots. \end{array}$$

Je dis que l'ensemble de tous ces éléments est dénombrable; il

---

[1] On peut remarquer que la propriété d'avoir même puissance que certaines de leurs parties aliquotes *caractérise* les ensembles infinis; cette propriété a pu être proposée comme définition de ces ensembles, par opposition aux ensembles finis.

suffit pour nous en convaincre de poser

$$u_i = s_{3i-2}, \qquad v_i = s_{3i-1}, \qquad w_i = s_{3i},$$

ou, si l'on préfère, d'écrire comme il suit les éléments les uns à la suite des autres, en inscrivant au-dessous de chacun d'eux le nombre entier positif correspondant, qui n'est autre que son *rang*

$$\begin{array}{cccccccccc} u_1, & v_1, & w_1, & u_2, & v_2, & w_2, & u_3, & v_3, & w_3, & \ldots, \\ 1, & 2, & 3, & 4, & 5, & 6, & 7, & 8, & 9, & \ldots. \end{array}$$

On conclut de là immédiatement qu'un ensemble dont tous les éléments ont pour indices des nombres entiers différents, positifs ou négatifs, est dénombrable.

Considérons maintenant un ensemble dont chaque élément a plusieurs indices, chaque indice prenant toutes les valeurs entières positives. Si nous nous bornons d'abord au cas de deux indices, nous pourrons écrire cet ensemble de la manière suivante :

$$\begin{array}{ccccc} u_{11} & u_{12} & u_{13} & u_{14} & \ldots. \\ u_{21} & u_{22} & u_{23} & u_{24} & \ldots, \\ u_{31} & u_{32} & u_{33} & u_{34} & \ldots, \\ u_{41} & u_{42} & u_{43} & \ldots & \ldots, \\ \ldots & \ldots & \ldots & \ldots & \ldots. \end{array}$$

Pour ranger ses éléments de telle manière qu'il apparaisse nettement qu'il est dénombrable, il nous suffira de les écrire comme il suit :

$$u_{11}; \quad u_{21}, \quad u_{12}; \quad u_{31}, \quad u_{22}, \quad u_{13}; \quad u_{41}, \quad u_{32}, \quad u_{23}, \quad u_{14}; \quad u_{51}, \quad \ldots.$$

Le procédé que nous venons d'employer s'étend immédiatement au cas où le nombre des indices dépasse deux ; considérons l'ensemble de tous les éléments

$$u_{m_1, m_2, \ldots, m_p},$$

$m_1, m_2, \ldots, m_p$ prenant toutes les valeurs entières et positives. Il est clair qu'il y a dans cet ensemble un nombre limité d'éléments pour lesquels on a

$$m_1 + m_2 + \ldots + m_p = n,$$

n étant un nombre positif déterminé. Si l'on suppose successive-

ment $n = p$, $p+1$, $p+2$, ..., on aura chaque fois un nombre limité d'éléments, que l'on pourra écrire sur une même ligne les uns à la suite des autres. On écrira d'abord l'élément $u_{1,1,\ldots,1}$ qui correspond à $n = p$, puis ceux qui correspondent à $n = p+1$ (en les rangeant dans un ordre quelconque), puis ceux qui correspondent à $n = p+2$, ... Comme, pour chaque valeur de n, on en écrit un nombre limité, il est clair que tout élément de l'ensemble occupera dans cette suite un rang fini, car il est obtenu pour une valeur finie de n; l'ensemble est donc dénombrable. Dans le cas où p est égal à 2 ou à 3, on peut donner à ce raisonnement une forme géométrique, que nous allons exposer, en supposant $p = 3$. Elle s'étendrait d'ailleurs aux valeurs de p supérieures à 3 par l'emploi de la géométrie à plus de trois dimensions.

Ayant choisi trois axes rectangulaires et une unité de longueur arbitraire, nous pouvons faire correspondre à l'élément u_{m_1, m_2, m_3} le point dont les coordonnées sont m_1, m_2, m_3. Nous obtenons ainsi tous les points dont les trois coordonnées sont des nombres entiers positifs. L'ensemble de ces points a manifestement même puissance que l'ensemble donné, d'après la manière même dont on l'a obtenu.

Or, cet ensemble de points a évidemment les propriétés suivantes : 1° il contient une infinité de points; 2° une sphère quelconque en renferme un nombre limité. Nous allons montrer que tout ensemble formé de points situés dans l'espace et possédant ces deux propriétés, a même puissance que E, c'est-à-dire est dénombrable. Pour le prouver, il nous suffit de faire voir que nous pouvons *numéroter* les points de cet ensemble au moyen des entiers positifs, de manière que deux points différents aient deux numéros différents et que chaque point ait un numéro unique.

Or, considérons une série de sphères concentriques de rayons croissants; nous supposerons, par exemple, que ces rayons sont successivement tous les nombres entiers positifs : la sphère S_n a pour rayon n. Par hypothèse, la sphère S_1 renferme un nombre limité (qui peut être nul) de points de l'ensemble : nous leur donnerons dans un ordre arbitraire les *numéros* 1, 2, 3, ..., p. La sphère S_2 en renferme aussi un nombre limité; considérons seulement ceux qui ne sont pas encore numérotés et donnons-leur les numéros $p+1$, $p+2$, ..., q; nous considérerons ensuite les

points (en nombre limité) situés à l'intérieur de S_3 et non encore numérotés; ils recevront les numéros $q+1$, $q+2$, ..., r. Il est clair que, en procédant ainsi, chaque point aura un numéro différent; d'ailleurs tous les points seront numérotés, car, si R est la distance de l'un d'eux P au centre commun des sphères et si $n-1$ désigne la partie entière de R, le point P sera numéroté lorsqu'on considérera la sphère S_n, ce qui arrivera certainement. L'ensemble des points considérés est donc dénombrable.

On voit que cette démonstration s'applique au cas où tous les indices ou quelques-uns d'entre eux prendraient aussi des valeurs entières négatives et aussi au cas où certains indices ne prendraient pas *toutes* les valeurs entières. Mais ce que nous avons dit au sujet du cas où il y a un seul indice doit suffire pour que le lecteur soit convaincu que ce sont là des détails sans importance ne pouvant introduire aucune difficulté. Nous ajouterons seulement une remarque : dans le cas où il y a plusieurs indices, il suffit que l'*un d'eux* puisse prendre une infinité de valeurs, le nombre des valeurs des autres pouvant être limité.

L'une des applications les plus importantes des considérations précédentes est relative aux nombres rationnels. On obtient tous les nombres rationnels $\frac{p}{q}$ en prenant pour q un nombre positif quelconque différent de zéro et pour p un nombre positif ou négatif premier avec q. Il est clair qu'on obtient ainsi tous les nombres rationnels, et chacun une seule fois. Or, on peut poser

$$\frac{p}{q} = u_{p,q},$$

p et q prenant les valeurs que nous venons de dire; l'ensemble des $u_{p,q}$ est manifestement dénombrable. De même, si nous considérons, dans un espace, a un nombre quelconque (*fini*) de dimensions, tous les points dont les coordonnées sont des nombres rationnels, ils forment aussi un ensemble dénombrable; par exemple, dans l'espace ordinaire, on peut faire correspondre au point

$$x = \frac{p}{q}, \qquad y = \frac{p'}{q'}, \qquad z = \frac{p''}{q''}$$

l'élément à six indices

$$u_{p,q,p',q',p'',q''}.$$

Ces six indices prennent d'ailleurs une *partie* des valeurs entières (positives ou négatives); cela suffit pour que l'ensemble des éléments considérés forme un ensemble dénombrable.

Remarquons maintenant que les observations présentées plus haut, sur la possibilité d'admettre que, dans une série à un seul indice, l'indice peut prendre seulement *une partie* des valeurs entières positives, sans que l'ensemble cesse d'être dénombrable, équivalent au théorème suivant : *si, par un procédé quelconque, on choisit une infinité d'éléments parmi les éléments d'un ensemble dénombrable, on obtient encore un ensemble dénombrable.*

Il en résulte que, si l'on considère, par exemple dans l'espace à trois dimensions, *une infinité de points à coordonnées toutes rationnelles*, on obtient un ensemble dénombrable, quel que soit le procédé par lequel on a distingué cette infinité de l'ensemble de tous les points à coordonnées toutes rationnelles.

Remarquons, enfin, que lorsque nous parlons des points à *coordonnées rationnelles*, nous devons toujours supposer que nous avons choisi une unité de longueur, laquelle est d'ailleurs arbitraire. On en conclut immédiatement que l'ensemble des points dont les coordonnées sont de la forme

$$x = \frac{m}{n}\alpha, \quad y = \frac{p}{q}\beta, \quad z = \frac{r}{s}\gamma,$$

m, n, p, q, r, s étant des entiers arbitraires et α, β, γ des *longueurs quelconques déterminées*, forme un ensemble dénombrable. Plus généralement, il en est de même des points dont les coordonnées peuvent se mettre sous la forme

$$\begin{aligned} x &= \frac{m}{n}\alpha + \frac{m'}{n'}\alpha' + \ldots + \frac{m^{(k)}}{n^{(k)}}\alpha^{(k)}, \\ y &= \frac{p}{q}\beta + \frac{p'}{q'}\beta' + \ldots + \frac{p^{(k)}}{q^{(k)}}\beta^{(k)}, \\ z &= \frac{r}{s}\gamma + \frac{r'}{s'}\gamma' + \ldots + \frac{r^{(k)}}{s^{(k)}}\gamma^{(k)}, \end{aligned}$$

les longueurs *déterminées* $\alpha, \ldots, \alpha^{(k)}, \beta, \ldots, \beta^{(k)}, \gamma, \ldots, \gamma^{(k)}$ pouvant être en nombre *quelconque*, mais *fini*, et les *entiers* m, $n, p, q, r, s, m', n', \ldots$ pouvant prendre toutes les valeurs possibles ou étant soumis à des restrictions de nature quelconque,

à condition, bien entendu, qu'il y ait une infinité de points dans l'ensemble.

Les exemples que nous avons donnés nous paraissent suffisants ([1]) pour faire comprendre au lecteur la nature des ensembles dénombrables, considérés en eux-mêmes ; nous allons maintenant les étudier dans leurs rapports avec les autres ensembles. Il est vrai que nous n'avons pas encore démontré l'existence d'ensembles non dénombrables ; mais nous pouvons provisoirement postuler cette existence, sans faire aucun cercle vicieux ; cette lacune ne tardera d'ailleurs pas à être comblée.

IV. — Comparaison des ensembles dénombrables avec les autres ensembles.

Nous allons d'abord montrer que, étant donné un ensemble *infini quelconque*, on peut toujours supprimer, parmi ses éléments, une infinité d'éléments formant un ensemble dénombrable, sans que l'ensemble cesse d'être infini. En désignant par A l'ensemble donné, nous pouvons écrire

$$A = D + A',$$

D désignant un ensemble dénombrable et A′ un ensemble infini ; le signe + signifie d'ailleurs que l'ensemble A est formé par la réunion des ensembles D et A′, *lesquels n'ont aucun élément commun*.

Pour démontrer cette proposition, il nous suffit de remarquer que, dire que l'ensemble A est infini, c'est dire que, étant donné un nombre quelconque $2n$, l'ensemble A renferme plus de $2n$ éléments. D'après cela, donnons-nous des nombres pairs croissants

$$2n_1, \quad 2n_2, \quad 2n_3, \quad 2n_4, \quad \ldots.$$

L'ensemble A renfermant au moins $2n_1$ éléments, nous pouvons ([2]) en choisir $2n_1$ et, parmi ces $2n_1$, en désigner un par u_1 ; l'en-

([1]) Dans le Chapitre suivant, nous étudierons un autre ensemble dénombrable très important, celui des nombres algébriques.

([2]) Sur la possibilité de tels choix, *voir* la Note IV, et en particulier la distinction entre les ensembles dénombrables et les ensembles effectivement énumérables. (*Note de la deuxième édition.*)

semble A renfermant au moins $2n_2$ éléments, nous pouvons en trouver $2(n_2 - n_1)$ *ne faisant pas partie des* $2n_1$ *déjà choisis* et désigner l'un d'eux par u_2, ... Il est clair que, en continuant ainsi, nous formerons un ensemble dénombrable

$$u_1, \quad u_2, \quad \ldots$$

que nous pouvons désigner par D et il subsistera une infinité A′ d'éléments, à savoir les $2n_1$ d'abord choisis, excepté u_1, les $2(n_2 - n_1)$ pris ensuite, sauf u_2, ..., plus ceux, s'il y en a (1), qui n'auront pas été choisis.

Ainsi on voit que tout ensemble infini comprend un ensemble dénombrable : en ce sens, on peut dire que la puissance des ensembles dénombrables est la plus petite des puissances.

Une proposition plus importante est la suivante : l'*ensemble infini* A′, que l'on obtient en supprimant de A les éléments de l'ensemble dénombrable D, *a même puissance que* A. En effet, de même que nous avons écrit

$$A = D + A',$$

nous pouvons écrire

$$A' = D' + A'',$$

D′ étant un ensemble dénombrable et A″ un ensemble infini. D'ailleurs, la réunion des deux ensembles dénombrables D et D′ forme un ensemble dénombrable Δ

$$D + D' = \Delta;$$

nous avons, par suite,

$$A = D + A' = D + D' + A'' = \Delta + A''.$$

Il s'agit de prouver que A et A′ ont même puissance ; il nous suffit de récrire les égalités

$$A = \Delta + A'',$$
$$A' = D' + A'',$$

et de remarquer que D′ et Δ ont même puissance, puisque ce sont deux ensembles dénombrables. A et A′ peuvent donc être

(1) On voit aisément qu'il y en aura si l'ensemble A n'est pas dénombrable. Dans ce cas il est évident que, si l'on supprime de A un ensemble dénombrable, il doit rester un ensemble non dénombrable, sinon A serait dénombrable.

regardés comme obtenus en ajoutant à un même ensemble A'' deux ensembles de même puissance D et Δ; ils ont donc évidemment même puissance [1].

Ce théorème est très important [2], parce qu'il nous montre que, lorsqu'on étudie les puissances des ensembles non dénombrables, on peut négliger sans inconvénient un ensemble dénombrable d'éléments. Les ensembles dénombrables jouent ainsi, par rapport aux autres, le même rôle que les infiniment petits vis-à-vis des quantités finies.

Nous allons montrer maintenant qu'il existe effectivement des ensembles non dénombrables en faisant voir que : *l'ensemble des nombres réels compris entre* o *et* 1 *n'est pas dénombrable*. Il suffit, pour le prouver, de montrer que, si l'on a un ensemble dénombrable quelconque

$$u_1, \quad u_2, \quad u_3, \quad \ldots, \quad u_n, \quad \ldots,$$

il y a des nombres compris entre o et 1 qui ne font pas partie de cet ensemble. Nous pouvons évidemment supposer que tous les u sont compris entre o et 1. Figurons le segment de droite o-1 et marquons-y les points u_1 et u_2, en supposant, pour fixer les idées, $u_1 < u_2$.

Considérons maintenant les éléments successifs de notre ensemble, dans leur ordre naturel,

$$u_3, \quad u_4, \quad u_5, \quad \ldots,$$

et soit u_{α_3} le premier d'entre eux qui est compris dans l'intervalle u_1-u_2 (on pourrait avoir $\alpha_3 = 3$). Il est clair que tous les points de l'ensemble compris dans l'intervalle u_{α_3}-u_2 ont un indice supérieur à α_3; car si l'un d'eux avait un indice inférieur à α_3 on aurait dû le prendre au lieu de u_{α_3}. Continuons à considérer les éléments successifs après u_{α_3} et soit u_{α_4} le premier

(1) Il résulte, en effet, de la définition même de la puissance que si deux ensembles peuvent être décomposés en un même nombre d'ensembles ayant respectivement mêmes puissances, ils ont même puissance.

(2) Comme exercice, le lecteur démontrera aisément d'une manière directe que l'ensemble des nombres réels, et l'ensemble A' des nombres réels *non entiers*, ont même puissance. Il pourra prendre pour D les nombres entiers n et pour D' les nombres de la forme $n + \frac{1}{2}$, par exemple.

d'entre eux qui est compris dans l'intervalle u_{α_3}-u_2; nous désignerons de même par u_{α_5} le premier élément compris entre u_{α_3} et u_{α_4} et, d'une manière générale, par $u_{\alpha_{n+2}}$ le premier élément compris dans l'intervalle u_{α_n}-$u_{\alpha_{n+1}}$. Il est clair d'abord que, si l'un de ces intervalles ne renfermait aucun point de l'ensemble, nous aurions constaté par cela même l'existence de points compris entre 0 et 1 et n'appartenant pas à l'ensemble; nous pouvons donc supposer que la suite des α_n se prolonge indéfiniment; d'ailleurs, les α_n étant les nombres entiers croissants, α_n augmente indéfiniment avec n. Enfin, d'après la manière même dont sont choisis les u_{α_n}, tous les éléments de l'ensemble compris entre u_{α_n} et $u_{\alpha_{n+1}}$ ont un indice supérieur à α_{n+1}. Cela posé, les nombres u_1, u_{α_3}, u_{α_5}, ..., $u_{\alpha_{2n+1}}$, ... sont visiblement croissants et tous inférieurs aux nombres décroissants u_2, u_{α_4}, u_{α_6}, ..., $u_{\alpha_{2n}}$, ... La première de ces suites a donc une limite v, la seconde une limite w, et l'on a $v \leqq w$. Je dis que le nombre v n'appartient pas à l'ensemble; en effet, si l'on avait

$$v = u_k,$$

nous pourrions prendre n assez grand pour que l'on ait $\alpha_n > k$ et nous arriverions à cette conclusion contradictoire que l'élément u_k est compris entre u_{α_n} et $u_{\alpha_{n+1}}$, k étant inférieur à α_n. Donc, il y a entre 0 et 1 au moins un point qui n'appartient pas à l'ensemble dénombrable *choisi arbitrairement*, u_1, u_2, ..., u_n, ... Il en résulte que l'ensemble des points compris entre 0 et 1 *n'est pas dénombrable*, et, par suite, qu'il y a une infinité non dénombrable de points qui n'appartiennent pas à l'ensemble dénombrable considéré (¹).

Nous connaissons donc un ensemble non dénombrable; on peut remarquer que l'existence de cet ensemble résulte du théorème suivant : *Des nombres croissants, tous inférieurs à un nombre fixe, ont une limite.* Or ce théorème est une conséquence de la *définition* des nombres incommensurables et, à quelque point de vue que l'on se place, postule la notion du *continu*. Nous ne nous

---

(¹) Ce raisonnement est *analogue* à celui-ci : Étant donné un ensemble, si nous savons démontrer que, lorsqu'on en supprime un nombre *fini quelconque* d'éléments, il en reste *au moins un*, nous pourrons affirmer que l'ensemble renferme une infinité d'éléments et, par suite, qu'il en reste une infinité.

étendrons pas plus sur l'origine et la nature de cette notion que nous ne l'avons fait sur la notion de l'ensemble E des nombres entiers positifs; nous admettrons que l'ensemble C des nombres compris entre 0 et 1 nous est *donné*, sans rechercher comment il pourrait l'être effectivement. Les ensembles ayant même puissance que l'ensemble C seront dits avoir *la puissance du continu*.

V. — Les ensembles qui ont la puissance du continu.

Parmi les ensembles ayant la puissance du continu, nous pouvons citer tout d'abord, d'après un théorème démontré tout à l'heure, les ensembles que l'on obtient en supprimant, parmi les nombres compris entre 0 et 1, *un ensemble dénombrable quelconque;* par exemple, en supprimant tous les nombres rationnels. D'autre part, il est aisé de voir que l'ensemble de tous les points situés sur une droite, ou sur une ou plusieurs courbes analytiques quelconques, a la puissance du continu, car on peut aisément établir une correspondance univoque et réciproque ([1]) entre les points de tels ensembles et les points compris entre 0 et 1.

Il est, d'autre part, évident que, si l'on réunit une infinité dénombrable d'ensembles dont chacun a la puissance du continu, l'ensemble ainsi obtenu a aussi la puissance du continu. Car on peut, par exemple, constater que le premier ensemble a même puissance que les points compris entre 0 et 1, le second même puissance que les points compris entre 1 et 2, le troisième même puissance que les points compris entre 2 et 3, etc., de sorte que l'ensemble formé par la réunion de l'infinité dénombrable d'ensembles considérée a même puissance que l'*ensemble de tous les nombres positifs*, c'est-à-dire la puissance du continu.

Mais un théorème bien plus remarquable et qui est une des plus belles découvertes de M. G. Cantor est le suivant : si l'on a une infinité d'ensembles ayant la puissance du continu, cette infinité (dans laquelle chaque ensemble est regardé comme un élé-

---

([1]) Si l'on cherche à établir analytiquement une telle correspondance elle pourra laisser de côté des points isolés en nombre limité : nous savons que cela n'a aucune importance.

ment) ayant elle-même la puissance du continu, l'ensemble total formé par la réunion de tous les éléments des ensembles considérés a aussi la puissance du continu. La forme géométrique que l'on peut donner à cet énoncé le rendra peut-être plus clair : considérons deux axes rectangulaires et un carré dont les côtés ont pour équations

$$x = 0, \quad x = 1; \quad y = 0, \quad y = 1.$$

Il est clair que l'on obtient tous les points intérieurs au carré par la réunion de tous les points situés sur toutes les parallèles au côté $y = 0$, menées par les divers points du côté $x = 0$. Sur chacune de ces parallèles on a un ensemble qui a la puissance du continu et l'ensemble de ces parallèles (chacune d'elles étant un élément) a évidemment aussi la puissance du continu. Le théorème de M. G. Cantor revient donc au suivant, et c'est sous cette forme qu'il a été énoncé par son auteur : *l'ensemble des points intérieurs au carré a même puissance que l'ensemble des points compris entre* 0 *et* 1.

Remarquons d'abord que nous pouvons nous borner à considérer les points à coordonnées incommensurables, car nous savons que l'ensemble Γ des nombres incommensurables compris entre 0 et 1 a même puissance que l'ensemble C de tous les nombres compris entre 0 et 1. Nous pouvons établir entre les éléments de Γ et de C une correspondance univoque et réciproque. Désignons par α, β, γ, ... les éléments de Γ qui correspondent respectivement aux éléments a, b, c, ... de C et faisons correspondre au point de coordonnées

$$x = a, \quad y = b,$$

le point de coordonnées

$$x = \alpha, \quad y = \beta;$$

nous constatons ainsi que l'ensemble des points intérieurs au carré (que l'on y comprenne ou non ceux qui sont sur le périmètre) a même puissance que l'ensemble des points intérieurs dont *les deux coordonnées sont incommensurables.*

Le théorème à démontrer revient donc à ceci : *l'ensemble des points intérieurs au carré et dont les deux coordonnées* x, y,

sont incommensurables, a même puissance que l'ensemble des nombres incommensurables z compris entre o *et* 1.

En d'autres termes, il suffit de faire voir que l'on peut, à tout système de deux nombres incommensurables x, y compris entre o et 1, faire correspondre un nombre incommensurable z compris entre o et 1, de manière qu'à tout couple x, y corresponde un seul nombre z et, à tout nombre z, un seul couple x, y.

Or il est très aisé de réaliser une pareille correspondance. Il nous suffit de remarquer avec M. G. Cantor que, si l'on réduit en fraction continue un nombre z compris entre o et 1, on obtient un développement de la forme

$$z = \cfrac{1}{\alpha_1 + \cfrac{1}{\alpha_2 + \cfrac{1}{\alpha_3 + \cfrac{1}{\alpha_4 + \ldots}}}},$$

$\alpha_1, \alpha_2, \alpha_3, \alpha_4, \ldots$ étant des entiers déterminés dont aucun n'est nul. Donc, à tout nombre z considéré correspond une suite bien déterminée d'entiers positifs, dont aucun n'est nul,

$$(1) \qquad \alpha_1, \quad \alpha_2, \quad \alpha_3, \quad \ldots,$$

et, réciproquement, une telle suite détermine d'une manière unique le nombre z. Cela posé, il nous suffira de déterminer x et y par deux suites infinies extraites de la suite (1), d'après une loi déterminée, par exemple par les deux suites

$$\alpha_1, \quad \alpha_3, \quad \alpha_5, \quad \ldots,$$
$$\alpha_2, \quad \alpha_4, \quad \alpha_6, \quad \ldots,$$

c'est-à-dire de poser

$$x = \cfrac{1}{\alpha_1 + \cfrac{1}{\alpha_3 + \cfrac{1}{\alpha_5 + \ldots}}}, \qquad y = \cfrac{1}{\alpha_2 + \cfrac{1}{\alpha_4 + \cfrac{1}{\alpha_6 + \ldots}}}$$

Il est clair que la connaissance de z détermine, d'une manière unique, x et y, et que, réciproquement, la connaissance de x et de y détermine z d'une manière unique. La proposition que nous avions en vue est donc établie. Il est clair que la démonstration

ne serait en rien modifiée, si, au lieu de deux variables x, y, on en avait un nombre fini quelconque.

Supposons maintenant que nous ayons une infinité (dénombrable) de variables

$$z_1, \quad z_2, \quad \ldots, \quad z_n, \quad \ldots$$

dont chacune peut prendre toutes les valeurs incommensurables comprises entre 0 et 1 et appelons *point* tout système de valeurs données à ces variables, deux points étant différents dès que *toutes* les variables n'ont pas les mêmes valeurs pour ces deux points. Je dis que l'ensemble des points ainsi obtenus a la puissance du continu. Pour le faire voir, posons, n étant un indice quelconque et les u des nombres entiers positifs,

$$z_n = \cfrac{1}{u_{n1} + \cfrac{1}{u_{n2} + \cfrac{1}{u_{n3} + \ldots + \cfrac{1}{u_{np} + \ldots}}}}$$

Nous savons, d'une infinité de manières, ranger les termes du tableau à double entrée

$$\begin{array}{lllll}
u_{11}, & u_{12}, & \ldots, & u_{1p}, & \ldots, \\
\ldots, & \ldots, & \ldots, & \ldots, & \ldots, \\
\ldots, & \ldots, & \ldots, & \ldots, & \ldots, \\
u_{n1}, & u_{n2}, & \ldots, & u_{np}, & \ldots, \\
\ldots, & \ldots, & \ldots, & \ldots, & \ldots,
\end{array}$$

dans une série simple :

$$v_1, \quad v_2, \quad \ldots, \quad v_k, \quad \ldots.$$

Choisissons un procédé déterminé (d'ailleurs quelconque) pour opérer cette transformation; nous poserons

$$z = \cfrac{1}{v_1 + \cfrac{1}{v_2 + \ldots}}$$

et nous ferons correspondre le point z au point $(z_1, z_2, \ldots, z_n, \ldots)$. Nous aurons ainsi réalisé la correspondance exigée; on lèverait aisément la restriction imposée aux z d'être incommensurables.

Les considérations précédentes sont très importantes parce qu'elles nous montrent que, *si l'on fait abstraction de la continuité de la correspondance entre deux ensembles continus*, il n'y a pas de différence essentielle entre les ensembles continus à une dimension et les ensembles continus à deux (ou trois ...) dimensions, entre les fonctions d'une variable et les fonctions de plusieurs variables. Nous entrerons plus loin dans quelques détails sur ce sujet (Note III); notre seul but, dans ce Chapitre, a été de préciser un peu la notion d'ensemble.

Voici la conclusion à laquelle nous arrivons; nous connaissons actuellement deux puissances : la puissance des ensembles denombrables et la puissance du continu; nous pouvons définir des ensembles ayant l'une de ces deux puissances en *supposant définis*, d'une part, l'ensemble E de tous les nombres entiers positifs, d'autre part, l'ensemble C de tous les nombres compris entre 0 et 1. Nous sommes d'ailleurs certains que les puissances de E et de C sont différentes. Ces notions nous suffiront pour les applications que nous avons en vue. Nous reviendrons dans les Notes sur plusieurs points que nous avons laissés de côté et chercherons notamment à éclaircir un peu les notions générales d'ensemble et de puissance.

CHAPITRE II.

LES NOMBRES ALGÉBRIQUES ET L'APPROXIMATION DES INCOMMENSURABLES.

I. — Généralités sur les nombres algébriques.

On appelle *nombre algébrique* toute racine d'une équation algébrique *à coefficients entiers* [1].

Soit $\alpha + i\beta$ une racine de l'équation

$$f(x) + ig(x) = 0, \tag{1}$$

f et g étant des polynomes à coefficients réels et entiers. Il est clair que l'équation

$$f(y) - ig(y) = 0 \tag{2}$$

admet la racine $\alpha - i\beta$.

Formons l'équation qui admet pour racines toutes les quantités de la forme

$$z = \frac{x+y}{2},$$

x étant une racine quelconque de (1) et y une racine quelconque de (2) et soit

$$F(z) = 0 \tag{3}$$

cette équation. On reconnaît immédiatement

1° Que cette équation a ses coefficients entiers [2];

2° Qu'elle a ses coefficients réels;

3° Qu'elle admet α pour racine.

---

[1] Il est clair que si les coefficients d'une équation sont *rationnels*, on peut les rendre entiers sans changer les racines de l'équation.

[2] Après que l'on a, s'il est nécessaire, multiplié son premier membre par un facteur convenable.

On verrait d'une manière analogue que β est racine d'une certaine équation algébrique à coefficients entiers réels.

Il en résulte qu'on peut se borner à considérer les nombres algébriques réels, c'est-à-dire les racines réelles des équations algébriques à coefficients entiers réels. Si α et β sont deux tels nombres, vérifiant respectivement les équations à coefficients entiers

$$F(\alpha) = 0, \tag{4}$$

$$G(\beta) = 0, \tag{5}$$

il est clair que $x = \alpha + i\beta$ vérifiera une équation à *coefficients entiers* que l'on obtiendra en éliminant β entre l'équation (5) et l'équation

$$F(x - i\beta) = 0. \tag{4'}$$

On obtiendra donc aisément tous les nombres algébriques, si l'on connaît les nombres algébriques réels. Cette proposition est d'ailleurs plus curieuse qu'utile et, si nous l'utilisons plus loin, on reconnaîtra aisément que nous pourrions nous en passer en modifiant à peine les raisonnements. Voici une proposition analogue, mais plus générale et plus importante : *les racines des équations à coefficients algébriques sont elles-mêmes des nombres algébriques.*

Considérons en effet une équation

$$f(x, \alpha, \beta, \ldots, \lambda) = 0, \tag{A}$$

f étant un polynome à coefficients entiers par rapport à l'inconnue x et aux nombres algébriques donnés α, β, ..., λ.

Par hypothèse, ces nombres sont des racines d'équations à coefficients entiers

$$\tag{B} \left\{ \begin{array}{l} \varphi(\alpha) = 0, \\ \psi(\beta) = 0, \\ \ldots\ldots\ldots, \\ \ldots\ldots\ldots, \\ \omega(\lambda) = 0. \end{array} \right.$$

Il est clair qu'en éliminant α, β, ..., λ entre les équations (A) et (B) on obtiendra une équation à coefficients entiers

$$F(x) = 0$$

qui admettra toutes les racines de (A); donc ces racines sont des nombres algébriques.

Il en résulte que *l'ensemble des nombres algébriques comprend tous les nombres que l'on peut définir à l'aide des nombres entiers supposés connus, et d'un nombre fini d'opérations algébriques.*

Cette proposition montre l'importance des nombres algébriques et fait prévoir qu'ils ont un grand nombre de propriétés arithmétiques intéressantes [1]; nous allons en indiquer quelques-unes.

II. — L'ensemble des nombres algébriques est dénombrable.

Parmi les équations à coefficients entiers que vérifie un nombre algébrique [2], il y en a une dont le degré est moins élevé que celui des autres [3]; ce degré sera dit le *degré du nombre algébrique.*

Considérons un nombre algébrique de degré n; il vérifie une équation de la forme

$$(a) \qquad m_0 x^n + m_1 x^{n-1} + \ldots + m_n = 0,$$

$m_0, m_1, \ldots, m_n$ étant des nombres entiers positifs ou négatifs [4].

---

[1] L'une des propriétés les plus importantes des nombres algébriques est la suivante : étant donné un nombre quelconque (fini) de nombres algébriques A, on peut déterminer un nombre algébrique ξ, tel que tous les nombres A soient égaux à des polynomes en ξ à coefficients entiers. Cette proposition se démontre aisément par les méthodes dues au génie de Galois; mais ce n'est point ici le lieu de les développer. On peut consulter à ce sujet quelques Chapitres de M. Jules Drach dans l'*Introduction à l'étude de la Théorie des nombres et de l'Algèbre supérieure* (Paris, Vuibert, 1895).

[2] Nous nous bornons, pour plus de netteté, aux nombres algébriques réels; il n'y aurait que très peu à modifier dans le cas contraire.

[3] Cette équation est *irréductible* dans le domaine naturel; mais nous n'avons ici nul besoin de la théorie de l'irréductibilité.

[4] Si nous tenons à préciser, c'est-à-dire à ne pas écrire deux fois la même équation, nous pouvons supposer : 1° que m_0 est positif, 2° que $m_0, m_1, \ldots, m_n$ sont *premiers entre eux dans leur ensemble.* Mais ces détails sont ici sans importance.

Si l'équation (α) a p racines réelles [1], nous les désignerons par

$$x_{m_0, m_1, m_2, \ldots, m_n, 1},$$
$$x_{m_0, m_1, m_2, \ldots, m_n, 2},$$
$$\ldots\ldots\ldots\ldots\ldots,$$
$$x_{m_0, m_1, m_2, \ldots, m_n, p},$$

dans un ordre d'ailleurs arbitraire.

Nous voyons ainsi que les nombres algébriques de degré n peuvent être représentés par une lettre x affectée de $n+2$ indices prenant seulement des valeurs entières (et ne les prenant d'ailleurs pas toutes, ce qui est sans importance); ils forment donc un ensemble dénombrable; si on les désigne par

$$y_1^{(n)},\ y_2^{(n)},\ y_3^{(n)},\ \ldots,\ y_q^{(n)},\ \ldots,$$

on voit que *tous* les nombres algébriques peuvent être représentés par une même lettre y affectée de deux indices entiers et *forment par suite un ensemble dénombrable*. Ce résultat fondamental est dû à M. G. Cantor qui l'a établi d'une manière un peu différente. Étant donnée une équation (α), appelons *rang* d'une racine quelconque de cette équation la somme

$$|m_0| + |m_1| + |m_2| + \ldots + |m_n| + n.$$

Il est manifeste que le nombre des équations de rang donné est limité. En considérant successivement les équations de rang 3, 4, 5, ..., on pourra ainsi ranger tous les nombres algébriques en une série simple, ce qui prouve que leur ensemble est dénombrable. Il est à peine besoin de faire observer que les nombres rationnels sont un cas particulier des nombres algébriques.

Une des conséquences importantes de la proposition de M. G. Cantor, c'est qu'il existe dans tout intervalle une infinité non dénombrable de nombres non algébriques. On peut même affirmer que l'ensemble des nombres non algébriques compris dans un intervalle donné, a la puissance du continu.

L'existence de nombres non algébriques a été reconnue pour la

---

(1) Si l'une de ces racines est un nombre algébrique de degré n, il en est de même des autres, d'après la théorie de l'irréductibilité. Si l'on n'admettait pas ce point, il suffirait de dire que l'on considère seulement celles des racines qui sont des nombres algébriques de degré n.

première fois par Liouville, à l'aide de considérations très intéressantes que nous allons développer dans un instant. Mais le procédé que nous indiquerons, d'après Liouville, pour former des nombres non algébriques est très artificiel et l'on ne connaît aucune propriété des nombres que l'on définit ainsi. Le problème qui consiste à déterminer si un nombre, défini par un procédé *analytique* déterminé, est ou n'est pas algébrique, est un des plus difficiles qui puissent se poser aux géomètres, car il semble qu'on ne puisse indiquer aucune voie générale pour l'aborder ; les moyens qui conduisent au but pourront varier complètement suivant la manière dont le nombre est défini.

Aussi, lorsque M. Hermite, en 1873, démontra que le nombre e n'est pas algébrique, ce résultat appela l'attention universelle. C'était en effet le premier exemple *effectif*, si l'on peut ainsi dire, d'un nombre *transcendant* (¹), c'est-à-dire le premier exemple d'un nombre transcendant défini d'une manière simple par l'analyse et non pas seulement par des séries arithmétiques, comme les nombres de Liouville. Cet exemple était d'ailleurs d'autant plus intéressant que le nombre e, au moins dans l'état actuel de la Science, est le plus important parmi ceux qui s'introduisent en Analyse.

D'ailleurs, une fois la voie ouverte par M. Hermite, on connut bientôt de nouveaux nombres transcendants. M. Lindemann, par une savante généralisation de la méthode employée par M. Hermite pour le nombre e, démontre que π est transcendant. Il démontre même la proposition plus générale suivante :

Si l'on a la relation

$$e^x = y,$$

les nombres x et y ne peuvent pas être tous les deux algébriques, à moins qu'on ait

$$x = 0, \qquad y = 1.$$

Si l'on pose $x = 1$, on a $y = e$; c'est le résultat de M. Hermite ; si $x = \pi i$, $y = -1$, on obtient le résultat de M. Lindemann cité en premier lieu.

(¹) Tout nombre non algébrique s'appelle *transcendant*.

Nous pouvons donc ajouter à la liste des nombres transcendants *les logarithmes naturels de tous les nombres algébriques* et *tous les nombres dont les logarithmes naturels sont des nombres algébriques.*

Ces propositions, dont l'importance est considérable, parce qu'elles nous font connaître des nombres transcendants très simplement définis, ne nous donnent, comme il est aisé de s'en assurer, qu'une infinité dénombrable de nombres transcendants. Au contraire, les procédés de Liouville, que nous allons maintenant exposer, en font connaître une infinité non dénombrable.

III. — Les recherches de Liouville.

La méthode de Liouville consiste essentiellement dans la considération des nombres rationnels qui *approchent* d'un nombre incommensurable déterminé.

Désignons par ξ un nombre algébrique réel *de degré n*, vérifiant l'équation à coefficients entiers with no rat

$$f(\xi) = 0.$$ poly. of equation

Soit $\frac{p}{q}$ une fraction irréductible, valeur *approchée* de ξ; d'une manière précise, nous supposons que $\frac{p}{q}$ est compris dans un certain intervalle α, β contenant ξ, et d'ailleurs quelconque; il est clair que, si x appartient à l'intervalle α, β, il existe un nombre M tel qu'on ait on (α, β)

$$(1) \qquad |f'(x)| < \mathrm{M}.$$ (α, β)

Cela étant, remarquons que, $f(x)$ étant un polynome à coefficients entiers, si l'on y remplace x par un nombre rationnel $\frac{p}{q}$, on obtiendra un résultat de la forme

$$f\left(\frac{p}{q}\right) = \frac{\mathrm{A}}{q^n},$$

A étant un nombre entier ; donc, si $\frac{p}{q}$ n'est pas racine [1] de l'équa-

---

[1] On peut toujours supposer que l'équation $f(x) = 0$ n'admet pas de racine rationnelle; sinon le nombre ξ vérifierait une équation plus simple.

tion $f(x)=0$, on a $|A| \geqq 1$ et, par suite,

$$(2) \qquad \left|f\left(\frac{p}{q}\right)\right| \geqq \frac{1}{q^n}.$$

Cela posé, supposons que $\frac{p}{q}$ soit dans l'intervalle α, β et remarquons que, par hypothèse, $f(\xi)$ est nul ; on a

$$f\left(\frac{p}{q}\right) = \left(\frac{p}{q} - \xi\right) f'\left[\xi + \theta\left(\frac{p}{q} - \xi\right)\right] \qquad 0 < \theta < 1.$$ MVT

On peut donc écrire, en vertu des inégalités (1) et (2),

$$M\left|\frac{p}{q} - \xi\right| > \frac{1}{q^n},$$

c'est-à-dire

$$(A) \qquad \left|\frac{p}{q} - \xi\right| > \frac{1}{Mq^n}.$$

Ainsi, étant donné un nombre algébrique ξ, de degré n, et un intervalle α, β comprenant ξ, *on peut déterminer un nombre* M *tel que, quel que soit le nombre* $\frac{p}{q}$ *dans l'intervalle* α, β, *on ait l'inégalité* (A).

Si nous supposons $q > M$, l'inégalité (A) devient

$$(B) \qquad \left|\frac{p}{q} - \xi\right| > \frac{1}{q^{n+1}}.$$

On voit d'ailleurs aisément que cette inégalité est, au moins pour les valeurs de q dépassant une certaine limite, vérifiée même lorsque $\frac{p}{q}$ est extérieur à l'intervalle α, β.

Nous pouvons donc affirmer que, ξ étant un nombre algébrique de degré n, *on a l'inégalité* (B) *dès que q dépasse une certaine limite* (limite qui ne dépend que de ξ).

Nous pouvons donc affirmer que, étant donné un nombre ξ, s'il existe des valeurs de q dépassant toute limite et telles qu'on ait pour chacune d'elles

$$(\beta) \qquad \left|\frac{p}{q} - \xi\right| < \frac{1}{q^{n+1}},$$

le nombre ξ ne peut pas être un nombre algébrique de degré n. Supposons maintenant que nous puissions faire la même démon-

stration pour chaque valeur de n, *nous pourrons affirmer que* ξ *est un nombre transcendant.*

Il importe de remarquer que, si l'on donne à $\frac{p}{q}$ et à ξ des valeurs déterminées, l'inégalité (β) ne peut être vérifiée que pour un nombre fini de valeurs de n; mais il peut arriver que, ξ étant donné, *quelle que soit la valeur fixe attribuée à* n, l'inégalité (β) soit vérifiée pour une infinité de valeurs de $\frac{p}{q}$; *mais ces valeurs ne restent pas toutes les mêmes lorsque* n *varie.* L'essentiel, pour qu'on puisse affirmer la transcendante de ξ, est qu'il y ait, pour chaque valeur de n, une infinité de valeurs de $\frac{p}{q}$ donnant lieu à l'inégalité (β).

Or, il est facile de former des nombres ξ ayant cette propriété; posons

$$(3) \qquad \xi = \frac{\alpha_1}{10} + \frac{\alpha_2}{10^{1.2}} + \frac{\alpha_3}{10^{1.2.3}} + \frac{\alpha_4}{10^{1.2.3.4}} + \ldots + \frac{\alpha_n}{10^{1.2.3\ldots n}} + \ldots,$$

$\alpha_1, \alpha_2, \ldots, \alpha_n, \ldots$ étant des nombres entiers positifs que, pour fixer les idées, je suppose inférieurs à 10, de sorte que ξ est en réalité écrit sous forme de fraction décimale ([1]). Désignons par $\frac{p}{q}$ la somme des m premiers termes de la série qui définit ξ; nous avons évidemment $q = 10^{1.2.3\ldots m}$ et

$$\xi = \frac{p}{q} + \frac{\alpha_{m+1}}{q^{m+1}} + \frac{\alpha_{m+2}}{q^{(m+1)(m+2)}} + \ldots.$$

On en conclut immédiatement

$$\left|\xi - \frac{p}{q}\right| < \frac{1}{q^m}.$$

Il suffit donc de prendre $m > n + 1$ pour prouver que l'inégalité (β), dans laquelle n a une valeur déterminée, est vérifiée pour une infinité de valeurs de $\frac{p}{q}$. Le nombre ξ est donc transcendant.

Il est aisé de voir qu'on peut ainsi définir une infinité non dénombrable de nombres transcendants et même, si l'on veut, une

---

([1]) On suppose qu'il y a une infinité de nombres α qui ne sont pas nuls.

infinité ayant la puissance du continu. Posons, en effet,

$$(4) \qquad \xi' = \frac{\alpha_1}{10} + \frac{\alpha_2}{10^2} + \frac{\alpha_3}{10^3} + \ldots + \frac{\alpha_n}{10^n} + \ldots$$

Si l'on donne aux α toutes les valeurs entières inférieures à 10, ξ' prendra toutes les valeurs comprises entre 0 et 1. Si l'on suppose de plus qu'il y a une infinité de nombres α qui ne sont pas nuls, on exclura en apparence certaines valeurs rationnelles des ξ'; mais on obtiendra ces valeurs en remarquant que 0,5, par exemple, peut s'écrire 0,49999 ... et que, dans la série (3), on n'a pas exclu le cas où, à partir d'un certain rang, tous les α sont égaux à 9. On a donc un ensemble ayant la puissance du continu. Or, si à chaque élément ξ' de cet ensemble, on fait correspondre le nombre ξ pour lequel les α ont la même valeur dans les formules (3) et (4), on aura défini un ensemble de nombres transcendants ξ ayant la puissance du continu. C'est le résultat que nous voulions obtenir.

IV. — L'approximation des nombres incommensurables.

La remarquable propriété des nombres algébriques, que nous avons démontrée d'après Liouville, nous amène à dire quelques mots de l'approximation des incommensurables en général. On sait que, étant donné un nombre incommensurable α, on peut trouver, d'une infinité de manières, une suite infinie de nombres commensurables ayant pour limite α; chacun d'eux peut être regardé comme une *valeur approchée* de α; mais il est clair que la connaissance d'une ou de plusieurs de ces *valeurs approchées* ne peut fournir aucun renseignement sur la nature arithmétique [1] de α; il en est tout autrement si l'on considère une *infinité* de valeurs rationnelles de plus en plus approchées, comme nous allons le voir. Mais il importe tout d'abord de mettre un peu d'ordre parmi ces valeurs approchées, et c'est là le but de la théorie des fractions continues arithmétiques. Je rappelle brièvement les résultats fondamentaux de cette théorie.

Étant donné un nombre incommensurable réel α, on peut le

---

[1] A cause de l'homogénéité du continu.

mettre, d'une seule manière, sous la forme

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ldots}}},$$

les a étant des nombres *entiers* dont le premier seul (a_0) peut être nul ou négatif, les autres étant essentiellement positifs. Si l'on pose

$$\begin{array}{llll} P_0 = 1, & P_1 = a_0, & \ldots, & P_{n+1} = a_n P_n + P_{n-1}, \\ Q_0 = 0, & Q_1 = 1, & \ldots, & Q_{n+1} = a_n Q_n + Q_{n-1}, \end{array}$$

les fractions *irréductibles* $\frac{P_n}{Q_n}$ sont les *réduites*.

Elles sont alternativement approchées par excès et par défaut, et l'on a

$$P_n Q_{n+1} - P_{n+1} Q_n = (-1)^n.$$

On déduit aisément de là la propriété fondamentale des réduites : chacune d'elles est plus approchée que toute fraction ayant les termes plus petits. En d'autres termes, $\frac{a}{b}$ étant une fraction quelconque et $\frac{P_n}{Q_n}$ une réduite, si l'on a

$$\left|\frac{a}{b} - \alpha\right| < \left|\frac{P_n}{Q_n} - \alpha\right|,$$

on peut en conclure $a > P_n$, $b > Q_n$.

Il résulte de là que, dans l'étude de l'approximation du nombre α au moyen de nombres rationnels, on peut en général ([1]) se borner à considérer les réduites. Il est d'ailleurs aisé d'indiquer deux limites de l'approximation obtenue à l'aide d'une réduite donnée. Nous supposons, pour plus de netteté, α positif.

Nous partirons des inégalités évidentes

$$\left|\frac{P_n}{Q_n} - \frac{P_{n+2}}{Q_{n+2}}\right| < \left|\frac{P_n}{Q_n} - \alpha\right| < \left|\frac{P_n}{Q_n} - \frac{P_{n+1}}{Q_{n+1}}\right|.$$

([1]) Dans certaines questions, il pourrait y avoir intérêt à considérer aussi les *fractions intermédiaires*. (*Voir*, par exemple, SERRET, *Algèbre supérieure*, 5e édition, t. I.)

En remarquant que

$$|P_n Q_{n+2} - P_{n+2} Q_n| \geqq 1$$

et

$$|P_n Q_{n+1} - Q_n P_{n+1}| = 1,$$

on en conclut

$$\text{(1)} \qquad \frac{1}{Q_n Q_{n+2}} < \left| \frac{P_n}{Q_n} - \alpha \right| < \frac{1}{Q_n Q_{n+1}}.$$

On a d'ailleurs

$$Q_{n+1} = a_n \quad Q_n \quad + Q_{n-1},$$
$$Q_{n+2} = a_{n+1} Q_{n+1} + Q_n$$

et l'on en déduit

$$a_n Q_n < Q_{n+1} < (a_n + 1) Q_n,$$
$$Q_{n+2} < (a_n a_{n+1} + a_{n+1} + 1) Q_n.$$

Les inégalités (1) deviennent enfin

$$\text{(2)} \qquad \frac{1}{(a_n a_{n+1} + a_{n+1} + 1) Q_n^2} < \left| \frac{P_n}{Q_n} - \alpha \right| < \frac{1}{a_n Q_n^2}$$

et fournissent deux limites de l'approximation donnée par la réduite de rang n.

Pour en donner une application immédiate, supposons que α soit un nombre algébrique déterminé de degré r, nous savons qu'il existe un nombre M tel qu'on ait, quels que soient p et q,

$$\left| \frac{p}{q} - \xi \right| < \frac{1}{M q^r}.$$

En prenant $p = P_n$, $q = Q_n$ et comparant avec la deuxième des inégalités (2), on en conclut

$$\text{(3)} \qquad a_n < M Q_n^{r-2}.$$

Ainsi, étant donné un nombre algébrique α de degré r, on peut trouver un nombre M tel que l'inégalité (3) ait lieu pour toute valeur de n. Dans le cas où $r = 2$, on voit que tous les a_n sont inférieurs à un nombre fixe M, ce qui s'accorde avec ce résultat bien connu que la fraction continue est alors périodique. Il est évidemment très aisé de former des fractions continues qui ne vérifient pas l'inégalité (3), soit pour une valeur déterminée de r, soit

même quel que soit r [1] : on a ainsi de nouveaux exemples de nombres transcendants.

Mais ces résultats, déduits de la seconde des inégalités (2), ne diffèrent pas sensiblement de ceux que nous avions obtenus par la simple considération des fractions décimales; au contraire, la première des inégalités (2), que nous récrirons sous la forme suivante

$$(4) \qquad \left|\frac{P_n}{Q_n} - \alpha\right| > \frac{1}{(a_n a_{n+1} + a_{n+1} + 1)Q_n^2},$$

va nous permettre d'obtenir des résultats auxquels ne conduirait pas la seule théorie des fractions décimales. Supposons que tous les quotients incomplets a_n soient compris entre 1 et 10, par exemple :

$$(5) \qquad 1 \leqq a_n < 10 \qquad (n = 1, 2, 3, \ldots),$$

et désignons par b un nombre entier quelconque compris entre Q_{n-1} et Q_n; nous avons

$$(6) \qquad Q_n > b > Q_{n-1} > \frac{1}{10} Q_n,$$

puisque $Q_n < (a_{n-1}+1)\,Q_{n-1}$ et $a_{n-1}+1 \leqq 10$. D'autre part, a étant un entier quelconque, on a, en utilisant les inégalités (4), (5), (6) et la propriété fondamentale des réduites,

$$\left|\frac{a}{b} - \alpha\right| > \left|\frac{P_n}{Q_n} - \alpha\right| > \frac{1}{100\,Q_n^2} > \frac{1}{10\,000\,b^2}.$$

Ainsi, l'inégalité (5) a pour conséquence la suivante

$$(7) \qquad \left|\frac{a}{b} - \alpha\right| > \frac{1}{10\,000\,b^2},$$

a et b étant des entiers arbitraires. Nous voyons ainsi que la propriété des nombres algébriques, due à Liouville, n'est nullement caractéristique et qu'on peut trouver une infinité de nombres transcendants α vérifiant l'inégalité (7), c'est-à-dire se comportant, au point de vue de l'approximation, comme les nombres algébriques du *second* degré.

---

[1] Il suffit, par exemple, de prendre $a_n > Q_n^n$.

Il est clair, en effet, que l'ensemble des nombres α, dont le développement en fraction continue satisfait aux inégalités (5), a même puissance que l'ensemble des nombres α' définis par l'égalité

$$\alpha' = \frac{a_1 - 1}{9} + \frac{a_2 - 1}{9^2} + \ldots + \frac{a_n - 1}{9^n} + \ldots,$$

les a vérifiant toujours les inégalités (5), et il est clair que l'ensemble des α' n'est autre [1] que l'ensemble des nombres compris entre 0 et 1 (y compris 0 et 1). L'ensemble des α a donc la puissance du continu et renferme par suite des nombres transcendants.

---

[1] *Voir* la remarque faite page 29 sur le cas où, à partir d'un certain rang, tous les a_n seraient égaux à 9 (et, par suite, les $a_n - 1$ à 8). D'ailleurs cette difficulté, ne concernant qu'une infinité dénombrable de points, est ici sans importance.

CHAPITRE III.

LES ENSEMBLES PARFAITS ET LES ENSEMBLES MESURABLES.

Nous avons considéré jusqu'ici les ensembles à un point de vue purement abstrait ; la notion de puissance peut être conçue indépendamment de la nature des éléments et une représentation concrète nous a été commode, mais non indispensable (¹) pour distinguer entre elles les puissances que nous avons appris à connaître. Nous allons, au contraire, étudier, dans ce Chapitre, deux notions de la plus grande importance, mais en relation étroite avec la signification concrète particulière des éléments de l'ensemble.

Nous supposerons que les éléments sont des points d'une droite, et le lecteur verra sans peine quelles considérations s'étendent au plan, ou à un espace d'un nombre quelconque de dimensions (²).

I. — Les ensembles fermés et les ensembles parfaits.

La notion fondamentale d'où découlent toutes les considérations développées dans ce paragraphe est celle d'*ensemble dérivé*. *On appelle ensemble dérivé* d'*un ensemble donné l'ensemble des points tels que*, *dans le voisinage de chacun d'eux*, *se trouve une infinité de points de l'ensemble donné*. En d'autres termes, soit A l'ensemble donné, la lettre a désignant l'un quelconque de ses éléments. On dira qu'un point a' est un *point*

(¹) On peut dire tout au moins qu'une représentation concrète *déterminée* n'est pas indispensable, ou, si l'on veut, que les propriétés étudiées (par exemple, p. 13) sont indépendantes de l'exemple concret choisi.

(²) On peut voir à ce sujet JORDAN, *Cours d'Analyse*, t. I, 2ᵉ édition.

limite de l'ensemble A si, quelque petit que soit ε, il existe un point a distinct de a' et dont la distance à a' soit inférieure à ε. L'ensemble des points limites a' constitue un ensemble A′, qui est dit *dérivé* de A. On voit que cette définition fait intervenir des éléments étrangers à l'ensemble lui-même, mais dépendant seulement de sa représentation concrète : par exemple, les points de la droite qui n'appartiennent pas à l'ensemble.

Pour donner immédiatement une application de cette définition, démontrons que, si l'ensemble A′ se réduit à un point, l'ensemble A est dénombrable. Remarquons d'abord qu'un intervalle fini qui ne renferme aucun point de A′ ne peut renfermer qu'un nombre limité de points de A [1]; dès lors, si nous supposons que le point unique de A′ a une abscisse nulle et si nous considérons tous les intervalles de l'une des deux formes

$$\left(\frac{1}{n}, \frac{1}{n+1}\right) \qquad (n, n+1),$$

n étant un nombre entier positif ou négatif (différent de 0 et de -1), chacun de ces intervalles renfermera un nombre limité de points de A. Or ces intervalles sont en infinité dénombrable et tout point de A est compris dans l'un d'eux (sauf le point zéro, s'il appartient à A) ; donc l'ensemble A est dénombrable.

On appelle ensemble parfait tout ensemble qui est identique avec son dérivé. Un exemple simple fera comprendre le sens de cette dénomination. Considérons l'ensemble A formé des points compris entre *zéro* et *un*, *non compris zéro* et *un*. Il est clair que l'ensemble A′ se compose de l'ensemble A, plus les points *zéro* et *un :* l'ensemble n'est pas parfait ; il le devient si on lui adjoint les points *zéro* et *un*. De même, l'ensemble des points intérieurs à un cercle n'est pas *parfait*, si l'on ne considère pas tous les points de la circonférence comme en faisant partie.

La définition que nous venons de donner est celle de M. G.

---

[1] Car s'il en renfermait une infinité, on pourrait le diviser en deux intervalles, chacun de ceux-là en deux, etc., et dans chaque division, un intervalle partiel au moins renfermerait une infinité de points; ces intervalles partiels, de plus en plus petits et compris les uns dans les autres, tendent vers un point limité.

Cantor; M. Jordan a donné une définition différente : il a appelé *ensemble parfait un ensemble qui contient tous les points de son dérivé*, mais qui peut contenir aussi des points n'appartenant pas à son dérivé. La dénomination d'*ensemble fermé* a prévalu pour les ensembles parfaits au sens de M. Jordan; on réserve, avec M. Cantor, le nom d'*ensembles parfaits* aux ensembles identiques à leur dérivé. La différence entre l'ensemble parfait et l'ensemble fermé est d'ailleurs moins grande qu'on ne pourrait le croire *a priori* en vertu du théorème suivant : *tout ensemble fermé ne diffère de son dérivé que par une infinité dénombrable de points.*

Il s'agit de prouver que, si un ensemble A contient tous les points de A′, l'ensemble B des points de A qui n'appartiennent pas à A′ est dénombrable. Supposons, pour fixer les idées, que tous les points de A appartiennent à un segment fini de droite; on verra aisément qu'une démonstration à peine différente s'appliquerait au cas où A est formé de points quelconques d'un espace à n dimensions. Soit b un point de B, c'est-à-dire un point de A qui n'appartient pas à A′; je dis qu'il existe un nombre h tel qu'il n'y ait pas de point de A′ dans l'intervalle $b-h$, $b+h$. Car, si un tel nombre h n'existait pas, il y aurait des points de A′ aussi rapprochés qu'on veut de b, et, par suite, des points de A aussi rapprochés qu'on veut de b, qui serait par suite un point de A′ contrairement à l'hypothèse [1].

Ce premier point établi, je dis que, h étant un nombre positif quelconque, les points de B tels qu'il n'existe pas de point de A′ dans l'intervalle $b-h$, $b+h$, *sont en nombre limité.* En effet, soit b_1 un de ces points; par hypothèse, l'intervalle b_1-h, b_1+h ne renferme pas de point de A′; il renferme donc un nombre limité de points de A et, par suite, un nombre limité de points de B. Soit b_2 un autre point de B, non situé dans l'intervalle b_1-h, b_1+h; l'intervalle b_2-h, b_2+h renfermera de même un nombre limité des points cherchés. Désignant par b_3 un des points cherchés,

---

[1] En d'autres termes, soit ε une constante positive quelconque; il existerait un point a' de A′ tel que $|a'-b|<\varepsilon$; mais, d'après la définition de A′, il existe un point a de A (distinct de b) et tel que $|a-a'|<\varepsilon$; on a donc $|a-b|<2\varepsilon$ et comme ε est arbitraire, b est un point de A′.

n'appartenant pas à ces deux intervalles, nous considérerons de même l'intervalle $b_3 - h$, $b_3 + h$.

Or, il est clair que, le point milieu de chacun de ces intervalles n'appartenant jamais à un autre intervalle, n de ces intervalles, dont chacun a une longueur $2h$, recouvrent, si l'on a soin de pas compter deux fois les parties communes, une longueur totale au moins égale à $(n+1)h$; leur nombre est donc forcément limité, et, par suite, le nombre des points de B ayant la propriété requise.

Considérons maintenant une suite arbitraire de nombres positifs décroissants et tendant vers zéro, par exemple :

$$1, \quad \frac{1}{2}, \quad \frac{1}{3}, \quad \ldots, \quad \frac{1}{n}, \quad \ldots,$$

et soit B_n l'ensemble des points de B définis en prenant $h = \frac{1}{n}$; B_n comprend un nombre limité de points; soit C_n l'ensemble des points qui appartiennent à B_n et n'appartiennent pas à B_{n-1} (il est clair que B_n renferme tous les points de B_{n-1}); on aura

$$B = C_1 + C_2 + C_3 + \ldots + C_n + \ldots$$

et, chacun des termes de cette somme renfermant un nombre limité de points, B est dénombrable. On remarquera que nous sommes assurés que tout point de B appartient à un groupe C_n de rang déterminé, par le théorème que nous avons démontré d'abord : à tout point b de B correspond un nombre fini h, tel qu'il n'y ait pas de point de A′ dans l'intervalle $b - h$, $b + h$.

On voit qu'un ensemble relativement parfait ne fait défaut à la définition de l'ensemble absolument parfait que par une infinité dénombrable de points. On ne peut cependant pas en conclure que *tout ensemble fermé est formé d'un ensemble parfait et d'un ensemble dénombrable*; car, si nous conservons nos notations, A est formé de l'ensemble A′ et de l'ensemble dénombrable B; mais l'ensemble A′ n'est peut-être pas parfait ; il coïncide avec le dérivé de A′, mais non nécessairement avec le dérivé de A′, car A′ n'est qu'une partie de A. La proposition que nous venons

d'énoncer est cependant exacte [1]; nous n'en aurons pas besoin [2].

On peut remarquer que, chemin faisant, nous avons implicitement démontré ce théorème : *tout ensemble dérivé* A′ *est relativement parfait*, c'est-à-dire contient son dérivé A″; car, tout point de A″, ayant dans son voisinage des points de A′, a aussi dans son voisinage des points de A, et par suite appartient à A′.

La notion opposée à celle d'ensemble parfait est celle d'ensemble *isolé* : on appelle ainsi tout ensemble qui n'a aucun élément commun avec l'ensemble dérivé; on démontre qu'un tel ensemble est dénombrable par un raisonnement identique à celui des pages 36, 37; mais, tout ensemble dénombrable n'est pas *isolé ;* par exemple, l'ensemble des nombres rationnels compris entre 0 et 1 admet comme ensemble dérivé l'ensemble de tous les nombres compris entre 0 et 1 ; il n'est donc pas isolé.

Nous dirons qu'un ensemble A est *dense* [3] *dans un intervalle* a, b, lorsque tout intervalle, quelque petit qu'il soit, renfermé dans a, b, contient des points de A. Dès lors, l'ensemble dérivé A′ comprend tous les points de a, b. Donc, un ensemble A, dense dans l'intervalle a, b, comprend, s'il est parfait, tous les points de a, b. Cette remarque fait comprendre l'importance de la notion d'*ensemble parfait ;* l'étude et la classification de ces ensembles est beaucoup plus simple que l'étude des ensembles non parfaits, surtout quand ceux-ci ne sont pas dénombrables. Mais j'ai hâte de quitter ces généralités pour montrer comment on se trouve naturellement conduit à considérer certains ensembles parfaits, lorsqu'on étudie de plus près la question qui a fait l'objet du Chapitre précédent.

---

(1) Le résultat essentiel d'où se déduit immédiatement cette proposition est dû à M. Bendixson : *Acta mathematica*, t. II, p. 419, théorèmes D, E, F.

(2) La démonstration suivante due à M. Lindelöf, peut s'exposer en quelques lignes : appelons *point de condensation* d'un ensemble, tout point dans le voisinage duquel se trouve une infinité *non dénombrable* de points de l'ensemble; les points de condensation d'un ensemble fermé appartiennent à cet ensemble; ils constituent un ensemble parfait; les autres points de l'ensemble fermé, c'est-à-dire ceux qui ne sont pas des points de condensation, forment un ensemble dénombrable; la décomposition est ainsi réalisée (*Note de la deuxième édition*).

(3) On a employé parfois l'expression : *condensé dans l'intervalle* a, b, dans le sens que nous donnons à *dense dans l'intervalle* a, b. Depuis la publication de la première édition de ce livre, l'expression *dense* que j'y proposais, paraît avoir prévalu.

II. — Les ensembles parfaits qui ne sont denses dans aucun intervalle.

Considérons les nombres rationnels compris entre o et 1 et associons à chacun d'eux $\frac{p}{q}$ l'intervalle

$$(\alpha) \qquad \frac{p}{q}-\frac{1}{q^3}, \quad \frac{p}{q}+\frac{1}{q^3}.$$

Nous obtenons ainsi une infinité dénombrable d'intervalles situés sur la droite; nous allons considérer l'ensemble A des points qui sont compris entre o et 1 et qui n'appartiennent à aucun de ces intervalles; on remarquera que chaque extrémité de tout intervalle, ayant une abscisse rationnelle, est le milieu d'un autre intervalle; il n'est donc pas nécessaire de spécifier si l'on considère les extrémités d'un intervalle comme en faisant ou n'en faisant pas partie : la définition de A n'en est pas modifiée.

Il est à peine besoin d'indiquer le lien étroit qu'il y a entre la considération de l'ensemble A et la question de l'approximation des incommensurables. Il est clair que A est formé des nombres incommensurables ξ ayant cette propriété qu'on a, quel que soit $\frac{p}{q}$,

$$\left|\xi-\frac{p}{q}\right|>\frac{1}{q^3}.$$

Qu'il existe de tels nombres ξ, nous en sommes assurés par la théorie des fractions continues; tel est, par exemple, le nombre $\frac{\sqrt{2}}{2}$; mais cette théorie nous apprend peu de chose sur l'ensemble A; de plus, elle repose sur des propriétés particulières des nombres rationnels et ne s'étendrait pas, au moins aisément, à l'approximation des incommensurables par des nombres algébriques d'une classe déterminée, par exemple.

Nous pouvons d'abord démontrer sans peine que l'ensemble A est parfait; car, si un point a' de A' n'appartenait pas à A, a' devrait être intérieur à l'un des intervalles (1) (sans coïncider avec ses extrémités), ce qui est absurde, puisque cet intervalle ne contient

pas de point de A. Donc, tout point de A′ appartient à A. Nous pouvons affirmer, par suite, que A n'est *dense* dans aucun intervalle ; car, si A était dense dans un intervalle, il comprendrait tous les points de cet intervalle, ce qui est absurde, puisque les points $\frac{p}{q}$ ne font pas partie de A et qu'il y en a dans tout intervalle.

Nous avons ainsi un exemple d'un ensemble fermé (¹), qui n'est dense dans aucun intervalle. Il importe d'approfondir la nature d'un tel ensemble et, tout d'abord, de démontrer qu'il n'est pas dénombrable, sans faire appel à la théorie des fractions continues.

Soit 0-1 un intervalle donné; sa *longueur* est égale à *un;* soit a_1-b_1 un intervalle compris dans l'intervalle 0-1 et de longueur α_1 ; il est clair que l'ensemble A des points qui appartiennent à l'intervalle 0-1 sans appartenir à l'intervalle a_1-b_1 est formé de tous les points de certains intervalles dont la longueur totale est $1-\alpha_1$. (Si, pour fixer les idées, on suppose $0<a_1<b_1<1$, ces intervalles sont 0-a_1 et b_1-1.)

Soit maintenant a_2-b_2 un autre intervalle compris dans l'intervalle 0-1 et n'ayant aucun point commun avec l'intervalle a_1-b_1; soit α_2 la longueur de a_2-b_2 (α_2, de même que α_1, est essentiellement positif). Il est clair que l'ensemble A des points de l'intervalle 0-1 qui n'appartiennent ni à a_1-b_1, ni à a_2-b_2 est formé de tous les points d'un nombre limité d'intervalles, dont la longueur totale est $1-\alpha_1-\alpha_2$.

Plus généralement, si l'on supprime de l'intervalle 0-1 un nombre limité d'intervalles a-b_1, a_2-b_2, ..., a_n-b_n, n'ayant aucune partie commune, et de longueurs respectives α_1, α_2, ..., α_n, l'ensemble A des points restants est formé de tous les points de certains intervalles, en nombre limité, et dont la longueur totale est

$$1-(\alpha_1+\alpha_2+\ldots+\alpha_n).$$

Si l'on ne faisait plus l'hypothèse que les ensembles a_1-b_1, a_2-b_2, ..., a_n-b_n n'ont pas de point commun, la conclusion

(¹) Nous savons d'ailleurs déjà par la théorie des fractions continues que A n'est pas dénombrable; nous pouvons donc, d'après le théorème de M. Bendixson, en déduire un ensemble *parfait* ayant la même propriété.

serait la même, sauf que l'on pourrait affirmer seulement que la somme des intervalles dont les points forment A est *supérieure ou égale* à $1 - (\alpha_1 + \alpha_2 + \ldots + \alpha_n)$. Dans tous les cas, si l'on suppose

$$\alpha_1 + \alpha_2 + \ldots + \alpha_n < 1,$$

cette somme n'est certainement pas nulle et, par suite, l'ensemble A a la puissance du continu.

Supposons maintenant que, à chaque nombre entier n, nous fassions correspondre un intervalle a_n-b_n, de longueur α_n; en d'autres termes, que nous supprimions de l'intervalle 0-1 les points d'une infinité dénombrable d'intervalles; supposons, en outre, que la série à termes positifs

$$s = \alpha_1 + \alpha_2 + \ldots + \alpha_n + \ldots$$

soit convergente et que sa somme s vérifie l'inégalité

$$s < 1;$$

que pourrons-nous dire de l'ensemble A formé des points de l'intervalle 0-1 qui n'appartiennent à aucun des intervalles a_n-b_n? On peut remarquer que cet ensemble peut n'être dense dans aucun intervalle; il est d'ailleurs aisé de voir qu'il comprend tous les points des intervalles dans lesquels il est dense; mais une première question se pose avant toutes celles-là : cet ensemble A existe-t-il? c'est-à-dire : peut-on conclure de l'inégalité

$$s < 1$$

qu'il y a des points n'appartenant à aucun des intervalles a_n-b_n? Bien que ce point soit à peu pres évident, il ne sera pas inutile de le démontrer en toute rigueur, car cela nous fournira l'occasion de faire plusieurs remarques importantes.

La première de ces remarques est la suivante : nous voulons démontrer qu'il existe des points non intérieurs à certains intervalles; ou, en d'autres termes, que l'hypothèse d'après laquelle tout point du segment 0-1 appartient à l'un de nos intervalles est absurde; *je dis qu'on peut ne considérer comme appartenant à un intervalle que les points intérieurs*, *à l'exclusion des*

extrémités. En effet agrandissons, à chacune des extrémités, chaque intervalle a_n-b_n d'une fraction ε de sa longueur ; c'est-à-dire prenons $b_n b'_n = a_n a'_n = \varepsilon a_n b_n$, les points a'_n et b'_n étant d'ailleurs extérieurs à l'intervalle a_n-b_n.

L'intervalle a'_n-b'_n a pour longueur $(1+2\varepsilon) a_n b_n$; la somme de tous les intervalles a'_1-b'_1, a'_2-b'_2, ... est donc $s' = (1+2\varepsilon)s$; mais, si l'on a

$$s < 1,$$

on pourra toujours choisir un nombre positif ε tel qu'on ait

$$(1+2\varepsilon)s < 1.$$

Or l'hypothèse que tout point de l'intervalle 0–1 serait compris dans un des intervalles a_n-b_n (sans exclure les extrémités) conduirait à ce résultat, que tout point est compris dans un intervalle a'_n-b'_n (en excluant les extrémités) ; nous allons montrer que cette dernière hypothèse est incompatible avec l'inégalité

$$s' < 1.$$

Nous allons démontrer pour cela le théorème suivant, dans l'énoncé duquel il est *expressément* entendu que les mots *intérieur à un intervalle* excluent les extrémités.

Si l'on a sur un segment limité de droite une infinité dénombrable d'intervalles partiels, tels que tout point de la droite soit intérieur à l'un au moins des intervalles, il existe [1] *un* NOMBRE LIMITÉ *d'intervalles choisis parmi les intervalles donnés et ayant la même propriété (tout point de la droite est intérieur à, au moins, l'un d'eux)*. Numérotons nos intervalles d'après une loi quelconque, mais déterminée ; je dis qu'il existe un nombre N, tel que tout point de la droite soit à l'intérieur d'un intervalle dont le rang ne dépasse pas N. En effet, nier l'existence du nombre N, c'est affirmer que, quel que soit le nombre donné n, il existe sur la droite un point tel que tous les intervalles qui le

---

(1) On trouvera dans ma Thèse une autre démonstration de ce théorème, démonstration qui donne un moyen au moins théorique de déterminer effectivement les intervalles en nombre limité dont il est question.

renferment ont un numéro supérieur à n. Il est clair d'ailleurs que, si l'on divise le segment de droite en deux segments égaux, l'un au moins de ces segments aura la même propriété ; car, si pour chacun ces segments il existait un nombre N, soient N′ et N″ ces deux nombres, il suffirait de prendre pour N le plus grand des deux. Si nous continuons à diviser le segment en deux parties égales et si nous conservons toujours le segment pour lequel il n'existe pas de nombre N (ou l'un deux, s'il y en a plusieurs), nous obtiendrons des segments de plus en plus petits, renfermés les uns dans les autres et ayant la propriété suivante : *quel que soit le nombre n, chacun d'eux contient au moins un point qui n'est renfermé dans aucun intervalle de rang inférieur à n.* Mais ces segments emboîtés les uns dans les autres et dont chacun est égal à la moitié du précédent ont un point limite α ; ce point α est, par hypothèse, *à l'intérieur* d'un intervalle de rang déterminé k, puisque nous avons supposé dénombrable l'ensemble de nos intervalles ; *les extrémités a_k, b_k de cet intervalle ne coïncident d'ailleurs pas avec* α (à cause du sens restreint que nous attachons au mot *intérieur*) ; donc, cet intervalle a_k-b_k comprend tout entier l'un des segments qui ont pour limite α, ce qui est absurde, puisque les points de ce segment seraient ainsi tous compris à l'intérieur de cet intervalle a_k-b_k dont le rang est un nombre fixe. L'existence du nombre N est donc établie.

Mais il est évident que si des intervalles, en nombre limité N, sont tels que tous les points d'un segment leur sont intérieurs, la somme des longueurs des intervalles est supérieure à la longueur du segment. C'est ce qui n'est pas possible, si les N intervalles sont choisis parmi une infinité d'intervalles, dont la somme *totale* est inférieure à la longueur du segment.

Nous pouvons donc affirmer que l'hypothèse

$$s' < 1;$$

ou, ce qui revient au même, l'hypothèse

$$s < 1$$

a pour conséquence l'existence certaine de l'ensemble A, c'est-à-dire l'existence de points n'appartenant pas aux intervalles

donnés (sans qu'il soit d'ailleurs maintenant nécessaire de distinguer si l'on exclut ou non les extrémités).

Nous pouvons ajouter que l'ensemble A n'est pas dénombrable, car s'il était dénombrable nous pourrions, désignant ses points par $\alpha_1, \alpha_2, \ldots, \alpha_n, \ldots$, entourer le point α_n d'un intervalle d'étendue égale à $\frac{\varepsilon}{2^n}$, et joindre ces intervalles aux intervalles donnés ; la somme s deviendrait

$$s_1 = s + \frac{\varepsilon}{2} + \frac{\varepsilon}{2^2} + \ldots + \frac{\varepsilon}{2^n} + \ldots = s + \varepsilon,$$

et il est clair que, si l'on a

$$s < 1,$$

on peut choisir ε de telle manière que l'on ait aussi

$$s_1 < 1.$$

Par conséquent, il existerait des points n'appartenant pas aux intervalles $a_n b_n$ et ne coïncidant certainement avec aucun des points $\alpha_1, \alpha_2, \ldots, \alpha_n, \ldots$. Il est donc impossible que cette suite dénombrable renferme tous les points de A; donc A n'est pas dénombrable.

Nous avons insisté sur cette démonstration parce qu'elle nous a paru de nature à éclairer un peu la conception que chacun peut essayer de se faire du *continu*. Après avoir réfléchi à ce fait, qu'on *peut enlever d'une droite tous les points compris dans chacun des intervalles*

$$\frac{p}{q} - \frac{1}{q^3}, \qquad \frac{p}{q} + \frac{1}{q^3},$$

et qu'il reste encore des points, en infinité non dénombrable, on sera moins disposé à croire qu'on sait ce que c'est que le continu et à raisonner sur lui comme sur une notion intuitive et parfaitement claire.

Voici un autre exemple, dans lequel on serait aisément amené à une conclusion inexacte.

Considérons les intervalles

$$\frac{p}{q} - \frac{\varepsilon}{q^3}, \qquad \frac{p}{q} + \frac{\varepsilon}{q^3} \qquad \begin{pmatrix} q = 2, 3, \ldots, & \infty \\ p = 1, 2, \ldots, q-1 \end{pmatrix}.$$

Leur étendue totale est, puisqu'il y en a $q-1$ correspondant à un même dénominateur q,

$$2\varepsilon\sum_{1}^{\infty}\frac{q-1}{q^3}=2\varepsilon\sum_{1}^{\infty}\left(\frac{1}{q^2}-\frac{1}{q^3}\right)=\varepsilon M,$$

en désignant par M un nombre aisé à calculer, inférieur à l'unité. D'ailleurs, si l'on tient à ne pas compter plusieurs fois leurs parties communes, on aura des intervalles, toujours en infinité dénombrable, et dont l'étendue totale sera inférieure à εM; mais ce point a peu d'importance.

Donnons à ε la valeur $\frac{1}{n}$, n étant un nombre entier; désignons par E_n l'ensemble formé de tous les points intérieurs aux intervalles correspondants, et considérons la suite des ensembles

$$(E)\qquad E_1,\quad E_2,\quad \ldots,\quad E_n,\quad \ldots;$$

il est clair que chacun d'eux renferme tous les points contenus dans les suivants; l'intervalle E_n est d'ailleurs formé de tous les points compris à l'intérieur d'intervalles dont la longueur totale est inférieure à $\frac{M}{n}$.

Ces préliminaires terminés, nous allons porter notre attention sur l'ensemble E formé des points qui appartiennent à *tous les* E_n; cet ensemble E comprend assurément tous les points $\frac{p}{q}$; mais il n'est pas évident *a priori* qu'il renferme d'autres points. Cet ensemble E a d'ailleurs la propriété remarquable qu'on peut enfermer ses points en une série d'intervalles dont la somme est aussi petite qu'on veut; car tous les points de E sont des points de E_n et, par suite, sont compris dans des intervalles d'une longueur totale inférieure à $\frac{M}{n}$. *Nous allons montrer que* E *a la puissance du continu.*

Il suffit [1] pour cela de faire voir que l'ensemble des nombres transcendants ξ étudiés page 28, appartient à l'ensemble E; c'est-à-dire de montrer que chacun de ces nombres ξ appartient à l'un

[1] *Voir* la Note I.

quelconque des ensembles E_n. Or chacun de ces nombres ξ est tel, qu'il existe une infinité de valeurs de p et de q vérifiant l'inégalité

$$\left|\xi - \frac{p}{q}\right| < \frac{1}{q^4}.$$

Puisqu'il y a une infinité de telles valeurs de q, nous pouvons supposer $q > n$ et nous aurons

$$\left|\xi - \frac{p}{q}\right| < \frac{1}{nq^3},$$

ce qui prouve que ξ appartient à l'ensemble E_n.

Nous sommes maintenant à même de comprendre une notion qui nous sera très utile, la notion d'ensemble mesurable [1].

III. — Les ensembles mesurables.

Tous les ensembles que nous considérerons seront formés de points compris entre 0 et 1. Lorsqu'un ensemble sera formé de tous les points compris dans une infinité dénombrable d'intervalles n'empiétant pas les uns sur les autres et ayant une longueur totale s, nous dirons que l'ensemble *a pour mesure* s. Lorsque deux ensembles n'ont pas de points communs, et que leurs mesures sont s et s', l'ensemble obtenu en les réunissant, c'est-à-dire leur somme, a pour mesure $s + s'$. D'ailleurs, il importe peu dans la définition de la mesure d'un ensemble, ou dans celle de la somme de deux ensembles, qu'on néglige ou qu'on tienne tel compte qu'on veut des extrémités des intervalles, en infinité dénombrable.

Plus généralement, si l'on a une infinité dénombrable d'ensembles n'ayant deux à deux aucun point commun et ayant respectivement pour mesures $s_1, s_2, \ldots, s_n, \ldots,$ leur somme (ou

[1] On comparera avec fruit les définitions que nous allons donner avec les définitions plus générales que donne M. Jordan dans son *Cours d'Analyse*. Le problème que nous étudions ici est d'ailleurs tout différent de celui qu'a résolu M. Jordan.

ensemble formé par leur réunion) a pour mesure

$$s_1 + s_2 + \ldots + s_n + \ldots$$

Tout cela est une conséquence de la définition de la mesure. Voici maintenant des définitions nouvelles : si un ensemble E a pour mesures s, et contient tous les points d'un ensemble E′ dont la mesure est s', *l'ensemble* E — E′, *forme des points de* E *qui n'appartiennent pas à* E′, *sera dit avoir pour mesure* $s - s'$; de plus, si un ensemble est la somme d'une infinité dénombrable d'ensembles sans partie commune, sa mesure sera la somme des mesures de ses parties et enfin les ensembles E et E′ ayant, en vertu de ces définitions, s et s' comme mesures, et E renfermant tous les points de E′, l'ensemble E — E′ aura pour mesure $s - s'$.

Le théorème fondamental démontré pages 41-43 nous assure que ces définitions ne seront jamais contradictoires entre elles [1] ; nous sommes donc libres de les adopter ; nous sommes d'ailleurs assurés aussi que la mesure d'un ensemble ne sera jamais une quantité négative ; mais un ensemble peut avoir pour mesure *zéro* et avoir la puissance du continu. Tel est l'ensemble E considéré tantôt. Si nous reprenons les notations de la page 45 et si nous désignons par α_n la mesure de E_n $\left(\alpha_n < \frac{M}{n}\right)$, l'ensemble $E_n - E_{n+1}$ aura pour mesure $\alpha_n - \alpha_{n+1}$ (nous savons que E_n renferme tous les points de E_{n+1}). L'ensemble A des points qui n'appartiennent pas a E_n a pour mesure $1 - \alpha_n$ (c'est la différence de l'ensemble de tous les points du segment 0-1 et de E_n.) L'ensemble des points qui n'appartiennent pas à E peut être regardé comme formé en ajoutant à A les ensembles $E_n - E_{n+1}$, $E_{n+1} - E_{n+2}$, ... ; sa mesure est donc

$$1 - \alpha_n + (\alpha_n - \alpha_{n+1}) + (\alpha_{n+1} - \alpha_{n+2}) + \ldots = 1,$$

puisque α_m tend vers zéro pour m infini. Donc, l'ensemble E, obtenu en retranchant cet ensemble de l'ensemble de tous les points 0-1, a pour mesure *zéro*.

Ainsi, un ensemble qui a pour mesure zéro peut être non

---

[1] Il est du moins aisé d'obtenir ce résultat par des procédés tout à fait analogues à ceux qu'on a employés pour établir ce théorème.

dénombrable; mais *tout ensemble dénombrable a pour mesure zéro;* c'est une conséquence aisée de ce qui précède.

Les ensembles dont on peut définir la mesure en vertu des définitions précédentes seront dits par nous mesurables, sans que nous entendions impliquer par là qu'il n'est pas possible de donner une définition de la mesure d'autres ensembles; mais une telle définition nous serait inutile; elle pourrait même nous gêner, si elle ne laissait pas à la *mesure* les propriétés fondamentales que nous lui avons attribuées dans les définitions que nous avons données ([1]).

Ces propriétés essentielles, que nous résumons ici parce qu'elles nous seront utiles, sont les suivantes : La mesure de la somme d'une infinité dénombrable d'ensembles est égale à la somme de leurs mesures; la somme de la différence de deux ensembles est égale à la différence de leurs mesures ([2]); *la mesure n'est jamais négative; tout ensemble dont la mesure n'est pas nulle n'est pas dénombrable.* C'est surtout de cette dernière propriété que nous ferons usage. Il est d'ailleurs expressément entendu que nous ne parlerons de mesure qu'à propos des ensembles que nous avons appelés *mesurables*.

Cependant, si un ensemble E contient tous les éléments d'un ensemble mesurable E_1, de mesure α, nous pourrons dire que la mesure de E est supérieure à α, sans nous inquiéter si E est mesurable ou non. Inversement, si E_1 contient tous les éléments de E, nous dirons que la mesure de E est inférieure à α. Les mots *supérieure* et *inférieure* n'excluent d'ailleurs pas l'égalité.

([1]) Le procédé que nous avons employé revient en réalité à ceci : nous avons reconnu qu'une définition de la mesure ne pouvait être utile que si elle avait certaines propriétés fondamentales : nous avons posé *a priori* ces propriétés et ce sont elles qui nous ont servi à définir la classe d'ensembles que nous regardons comme mesurables. Cette manière de procéder présente de grandes analogies avec les méthodes introduites par M. J. Drach, en Algèbre et dans la théorie des équations différentielles [*voir*, par exemple, l'Ouvrage cité (p. 23) et *Comptes rendus*, janvier 1895]. Dans tous les cas, elle procède à la même idée fondamentale : définir les éléments nouveaux qu'on introduit, à l'aide de leurs propriétés *essentielles*, c'est-à-dire de celles qui sont strictement indispensables pour les raisonnements qui doivent suivre.

([2]) Bien entendu, quand on parle de la somme de plusieurs ensembles, on suppose qu'ils n'ont, deux à deux, aucun élément commun et, quand on parle de leur différence, on suppose que l'un d'eux renferme tous les éléments de l'autre.

Il est aisé de voir que les propriétés essentielles s'étendent, avec des modifications convenables, à ces nouvelles définitions : en quelque sorte, un calcul d'égalités se trouve remplacé par un calcul d'inégalités qui peut parfois rendre les mêmes services (1).

Nous allons démontrer, en terminant, une proposition importante, qui montrera la liaison intime entre les notions diverses introduites dans ce Chapitre : *tout ensemble parfait* (2) *limité* (3) *est mesurable.*

Soit en effet A un ensemble parfait dont tous les points sont dans l'intervalle o-1 et soit α un point de cet intervalle, n'appartenant pas à A, je dis qu'il existe un intervalle ab comprenant α et ne renfermant pas de point de A, a et b étant d'ailleurs des points de A. (L'un des points a ou b pourrait ne pas appartenir à A; il coïnciderait alors avec le point o ou le point 1.) En effet, A étant parfait, α n'appartient pas à A'; il existe donc un nombre ε tel qu'il n'y ait pas de point de A dans l'intervalle α, $\alpha+\varepsilon$, ε étant positif. D'autre part, en excluant le cas où il n'y aurait pas de point de A dans l'intervalle α-1, il existe un nombre ε' tel qu'il y ait au moins un point de A dans l'intervalle α, $\alpha+\varepsilon'$. Il est manifeste que tout nombre positif η, ou bien a la même propriété que ε, ou bien a la même propriété que ε', et l'on voit aisément qu'il existe un nombre b tel que, si $\eta < b$, il n'y a pas de point de A dans l'intervalle α, $\alpha+\eta$, tandis qu'il y en a si $\eta > b$. Or, il est aisé de voir que le point b, ou bien appartient à A, ou bien (4) appartient à A'; or, A est parfait; donc, dans tous les cas, b appartient à A. On démontrerait de même l'existence d'un point a en prenant ε et η négatifs; l'intervalle ab est l'intervalle cherché.

Ainsi, A étant un ensemble parfait, tout point α qui n'appartient pas à A se trouve dans un intervalle ab ne renfermant d'autre point de A que ses extrémités. Si nous considérons maintenant

(1) *Voir* la Note VI (*Note de la deuxième édition*).

(2) Le théorème est vrai aussi pour les ensembles *fermés*.

(3) On dit qu'un ensemble est limité lorsque la distance de deux quelconques de ses points est inférieure à un nombre fixe.

(4) Car, si le point b n'appartient pas à A, il y a des points de A distincts de b, dans l'intervalle b, $b+h$, quelque petit que soit le nombre positif h; donc b appartient à A'

un autre point α', ne faisant pas partie de A et non situé dans ab, nous aurons un intervalle analogue $a'b'$, etc. D'ailleurs, l'ensemble des intervalles qu'on peut ainsi définir est certainement dénombrable, puisque la somme de leurs longueurs est au plus égale à l'unité. L'ensemble A s'obtient donc en retranchant de l'ensemble 0-1 une infinité dénombrable d'ensembles mesurables; il est donc mesurable [1].

Cette démonstration nous fait en même temps connaître le moyen le plus général de construction d'un ensemble parfait situé sur le segment 0, 1. Pour se donner un tel ensemble, il suffit de se donner une infinité dénombrable d'intervalles $a_1 b_1$, $a_2 b_2$, ..., $a_n b_n$, ..., assujettis d'ailleurs à certaines restrictions [2]. On en conclut aisément que l'ensemble de tous ces ensembles parfaits a la puissance du continu. Nous verrons que, d'autre part, l'ensemble de tous les ensembles de points situés sur le segment 0, 1, a une puissance plus élevée que celle du continu [3]. Cette remarque suffit à montrer quelle restriction considérable on apporte à la notion d'ensemble, lorsqu'on assujettit un ensemble à être fermé, ou *a fortiori* à être parfait et, par suite, combien il est naturel qu'on puisse être conduit ainsi à des résultats plus simples.

---

[1] Nous avons eu d'ailleurs un exemple d'un ensemble mesurable non fermé; car il était dense dans tout intervalle.

[2] Ces intervalles ont été appelés par M. Baire, intervalles *contigus* à l'ensemble parfait. (*Note de la deuxième édition.*)

[3] *Voir* Note I, pages 109-110.

CHAPITRE IV.

LE PROLONGEMENT ANALYTIQUE.

Les Chapitres précédents ont été surtout consacrés à la théorie des ensembles; ceux qui suivent contiendront les applications de cette théorie à la théorie des fonctions.

I. — La définition du prolongement analytique.

La notion de prolongement analytique est devenue trop classique pour qu'il soit nécessaire d'en exposer en détail les principes : nous allons nous borner à rappeler sans démonstration les résultats essentiels sur lesquels elle est fondée. Nous désignerons, suivant un usage assez répandu, par $\mathfrak{P}(x-a)$ une série ordonnée suivant les puissances positives de $x-a$ et convergente dans un certain cercle, dont le rayon est différent de zéro et dont le centre est le point a.

Cela posé, la première proposition fondamentale est la suivante : *Étant données deux séries $\mathfrak{P}(x-a)$ et $\mathfrak{P}(x-b)$, si leurs cercles de convergence ont une partie commune et si les deux séries ont même valeur, ainsi que toutes leurs dérivées en un point* INTÉRIEUR *à cette partie commune, elles ont même somme en tous les points où elles sont toutes deux convergentes*. Dès lors, chacune des deux séries est dite un *prolongement analytique* [1] *immédiat* de l'autre. En pratique, pour obtenir un prolongement analytique d'une série donnée $\mathfrak{P}(x-a)$, on est amené à prendre un point b, *intérieur* au cercle de convergence de $\mathfrak{P}(x-a)$ et à développer $\mathfrak{P}(x-a)$ suivant les puis-

[1] Il est à peine besoin de remarquer que, lorsqu'on donne $\mathfrak{P}(x-a)$ et b, $\mathfrak{P}(x-b)$, s'il existe, est déterminé d'une manière unique.

sances positives de $x-b$, par la formule de Taylor; mais il n'y a aucun inconvénient à employer l'expression *prolongement analytique immédiat* dans le sens un peu plus large que nous lui donnons; seulement, étant donnée une série $\mathfrak{P}(x-a)$, on n'a pas de procédé de calcul direct pour reconnaître s'il existe une série $\mathfrak{P}(x-b)$, ayant les propriétés voulues, lorsque b est extérieur au cercle de convergence de $\mathfrak{P}(x-a)$.

Voici maintenant la deuxième proposition fondamentale : *Si l'on a plusieurs séries $\mathfrak{P}(x-a)$, $\mathfrak{P}(x-b)$, $\mathfrak{P}(x-c)$, ..., $\mathfrak{P}(x-l)$, telles que leurs cercles de convergence recouvrent complètement une aire S à contour simple, et cela de telle manière que tout point de S soit* INTÉRIEUR *à l'un au moins des cercles; si, de plus, on peut ranger ces séries en une suite*

$$\mathfrak{P}(x-a),\quad \mathfrak{P}(x-b),\quad \ldots,\quad \mathfrak{P}(x-l),$$

les comprenant toutes, et telle que deux consécutives soient toujours le prolongement analytique immédiat l'une de l'autre, on pourra affirmer que, dans le cas où les cercles de convergence de deux séries non consécutives QUELCONQUES *se trouvent avoir une partie commune, ces deux séries sont le prolongement analytique immédiat l'une de l'autre.*

Dès lors, l'ensemble des séries considérées définit une fonction analytique uniforme dans S; chacune de ces séries est dite, d'après Weierstrass, un *élément* de cette fonction analytique.

Considérons maintenant une courbe fermée C sans point double; soient a, b, c, ..., l des points de C et $\mathfrak{P}(x-a)$ et $\mathfrak{P}(x-b)$, ..., $\mathfrak{P}(x-l)$ une suite de séries dont chacune est un prolongement analytique immédiat de la précédente; on suppose que, en parcourant la courbe dans un sens déterminé, on rencontre successivement les points a, b, c, ..., l. Supposons maintenant que les cercles de convergence de $\mathfrak{P}(x-l)$ et $\mathfrak{P}(x-a)$ aient une partie commune et que, d'autre part, tout point de la courbe soit intérieur à l'un au moins des cercles de convergence. L'ensemble de ces cercles recouvre ainsi une sorte de couronne à laquelle la courbe est intérieure.

Cela posé, deux cas généraux pourront se présenter :

1° La série $\mathfrak{P}(x-a)$ est un prolongement immédiat de $\mathfrak{P}(x-l)$; nos séries définissent alors une fonction *uniforme dans la cou-*

ronne, mais non nécessairement ([1]) uniforme à l'intérieur de C;

2° La série $\mathfrak{P}(x-a)$ n'est pas un prolongement analytique de $\mathfrak{P}(x-l)$; alors la fonction n'est pas uniforme dans la couronne C et l'on pourra, en général, tourner indéfiniment dans le même sens, en suivant C, sans retrouver une série qui soit le prolongement analytique immédiat de $\mathfrak{P}(x-a)$. Si cependant, après avoir décrit m fois le contour C, on trouve une telle série, la fonction est dite *algébroïde* à m déterminations dans la couronne; elle n'est d'ailleurs pas nécessairement ([2]) algébroïde à l'intérieur de C.

Donnons-nous maintenant, *a priori*, une série $\mathfrak{P}(x-a)$; si elle a des prolongements analytiques immédiats, nous les obtiendrons certainement, en prenant divers points à l'intérieur de son cercle de convergence et en formant le développement de Taylor correspondant; nous procéderons de même avec ces nouvelles séries et, en continuant indéfiniment, nous obtiendrons tous les éléments de fonction analytique qu'on peut déduire de l'élément donné et qui constituent avec lui une *fonction analytique*. Telle est la définition la plus générale donnée par Weierstrass de la fonction analytique. Nous y reviendrons plus loin (Chap. VI) pour la discuter et tenter de la généraliser; nous allons nous contenter ici d'en étudier avec quelque détail certaines conséquences.

II. — Un théorème de MM. Poincaré et Volterra.

Nous allons exposer d'abord une remarque importante qu paraît avoir été faite d'une manière indépendante par M. Poincaré et M. Volterra ([3]).

Étant donné un élément $\mathfrak{P}(x-a)$ de fonction analytique,

---

([1]) Il suffit de considérer la fonction $\sqrt{(z-\alpha)(z-\beta)}$ et un contour C entourant les deux points α et β.

([2]) Tel est le cas, par exemple, de la fonction $\sqrt{z}+\log\frac{z-\alpha}{z-\beta}$, si le contour C entoure les points α, β, o.

([3]) H. Poincaré, *Sur une propriété des fonctions analytiques* (*Rendiconti del Circolo matematico di Palermo*, t. II). — Volterra, *Sulle funzioni analitiche polidrome* (*Atti della Reale Accademia dei Lincei*. Serie quarta, *Rendiconti*, IV_2, p. 355).

il est clair qu'il existe une infinité non dénombrable d'éléments $\mathfrak{P}(x-b)$, dont chacun est un prolongement analytique immédiat ([1]) de $\mathfrak{P}(x-a)$. Chacun de ces éléments $\mathfrak{P}(x-b)$ permet de définir de même une infinité non dénombrable d'éléments $\mathfrak{P}(x-c)$, et ainsi de suite. On peut donc se demander d'abord s'il est possible d'indiquer un procédé régulier pour trouver sûrement la valeur de la fonction en un point qu'il est possible d'atteindre, et ensuite, dans le cas où la fonction n'est pas uniforme, si l'on est certain d'obtenir *toutes* les valeurs de la fonction en un point, en employant une méthode régulière. Il est aisé de se rendre compte que ces questions sont intimement liées avec la suivante : quelle est la puissance de l'ensemble des valeurs que peut prendre la fonction en un point donné? Il est clair, en effet, que si cette puissance est la *première*, c'est-à-dire si cet ensemble de valeurs est dénombrable, on pourra espérer les avoir toutes au moyen d'une suite régulière d'opérations, c'est-à-dire d'opérations effectuées successivement dans un ordre déterminé et formant par suite aussi un ensemble dénombrable. Au contraire, si la puissance de cet ensemble de valeurs dépassait la première, on ne pourrait pas être certain d'obtenir l'une quelconque d'entre elles après un nombre suffisamment grand d'opérations, quelle que soit d'ailleurs la loi à laquelle on assujettisse ces opérations.

Voici la réponse donnée par MM. Poincaré et Volterra à ces questions : *On peut définir toute fonction analytique au moyen d'une infinité dénombrable d'éléments* $\mathfrak{P}(x-a)$; par suite, l'ensemble des valeurs qu'une telle fonction prend en un point est dénombrable, et l'on obtiendra sûrement l'une quelconque de ces valeurs par un nombre *fini* d'opérations. Il est clair, d'autre part, que si le nombre de valeurs est infini, il faudra une infinité d'opérations pour les avoir toutes; mais on est assuré que, l'ordre des opérations à faire étant fixé une fois pour toutes, chaque valeur sera obtenue par une opération de rang déterminé.

Voici la remarque très simple à l'aide de laquelle MM. Poincaré et Volterra ont obtenu ces résultats essentiels ([2]) :

([1]) Le seul cas d'exception est celui où la série $\mathfrak{P}(x-a)$ n'admettrait *aucun* prolongement analytique; son cercle de convergence serait une coupure.

([2]) On rapprochera avec intérêt ces résultats du suivant, tout à fait analogue, à celui qu'énonce M. Weierstrass dans sa lettre à M. Schwarz (*Weierstrass*

Considérons un élément $\mathfrak{P}(x-a)$ et les cercles de convergence C de tous les éléments $\mathfrak{P}(x-b)$ qu'on obtient par la formule de Taylor, en prenant pour b un point quelconque *intérieur* au cercle de convergence de $\mathfrak{P}(x-a)$. Parmi les cercles C, désignons par C′ ceux dont le centre a pour coordonnées *deux nombres rationnels* : TOUT POINT INTÉRIEUR A L'UN DES CERCLES C EST INTÉRIEUR A L'UN DES CERCLES C′.

Or, nous savons que les cercles C′ forment un ensemble dénombrable; soient $b_1, b_2, \ldots, b_n, \ldots$ leurs centres; nous aurons une infinité dénombrable d'éléments $\mathfrak{P}(x-b_1)$, $\mathfrak{P}(x-b_2)$, ..., $\mathfrak{P}(x-b_n)$, ..., dont la considération, au point de vue de la définition de la fonction analytique, rendra *exactement* les mêmes services que l'infinité *non dénombrable* de tous les éléments $\mathfrak{P}(x-b)$. Nous pourrons, de même, à chaque élément $\mathfrak{P}(x-b_i)$, faire correspondre une infinité dénombrable d éléments

$$\mathfrak{P}(x-c_{i1}),\quad \mathfrak{P}(x-c_{i2}),\quad \ldots,\quad \mathfrak{P}(x-c_{in}),\ \ldots$$

et l'ensemble dénombrable des éléments

$$(\gamma)\qquad \mathfrak{P}(x-c_{ik})\qquad (i, k = 1, 2, \ldots, \infty)$$

rendra exactement les mêmes services que l'ensemble non dénombrable des éléments

$$(\gamma')\qquad \mathfrak{P}(x-c),$$

qu'on obtiendrait en prenant pour c un point quelconque intérieur au cercle de convergence de l'un des éléments $\mathfrak{P}(x-b)$.

En d'autres termes, tout point *intérieur* au cercle de convergence de l'un des éléments (γ) est intérieur au cercle de convergence de l'un des éléments (γ'). Si l'on continue de même, on verra que l'ensemble des éléments qu'on obtient en répétant de nouveau la même opération [c'est-à-dire en prenant tous les points à coordonnées rationnelles, intérieurs aux cercles de convergence des éléments (γ)], éléments qu'on peut désigner par

$$(\delta)\qquad \mathfrak{P}(x-d_{ikl})\qquad (i, k, l = 0, 1, 2, \ldots, \infty),$$

Werke, t. II, p. 236) : toute fonction algébrique peut être définie à l'aide d'un nombre *fini* d'éléments (en supposant que le point à l'infini est ordinaire et que $\mathfrak{P}\left(\frac{1}{x}\right)$ est la forme des éléments correspondants).

donne tous les points *intérieurs* aux cercles de convergence des éléments

$$(\delta') \qquad \mathfrak{P}(x-d),$$

qu'on obtiendrait en partant des éléments (γ') et en prenant pour d tous les points h intérieurs à leurs cercles de convergence.

On peut continuer de même et répéter cette opération une infinité dénombrable de fois; car, d'après la définition de Weierstrass, le domaine dans lequel est définie la fonction analytique s'obtient en répétant m fois l'opération qui nous a permis de passer des éléments (γ') aux éléments (δ') et en faisant croître m indéfiniment. La fonction est définie en tout point *intérieur* à l'un des cercles de convergence ainsi obtenus. et correspondant à une valeur finie de m. Or, nous savons que, si nous effectuons parallèlement l'opération qui nous a permis de déduire les éléments (δ) des éléments (γ), nous obtiendrons, après m opérations, un ensemble d'éléments qui pourra être représenté par

$$(\mu) \qquad \mathfrak{P}(x-l_{\alpha_1,\alpha_2,\ldots,\alpha_m}) \qquad (\alpha_1, \alpha_2, \ldots, \alpha_m = 1, 2, 3, \ldots, \infty)$$

et que le point A sera certainement intérieur à l'un des cercles de convergence de l'ensemble (μ).

Or, nous savons que l'ensemble des éléments tels que (μ), et correspondant à toutes les valeurs de m, est dénombrable, car c'est un ensemble dénombrable d'ensembles dénombrables.

Donc, *une infinité dénombrable d'opérations nous donnera la valeur de la fonction en tous les points où elle est définie d'après Weierstrass.* Par suite, si en un point elle a une infinité de valeurs, *l'ensemble de ces valeurs sera dénombrable.* C'est ce dernier résultat que se proposait d'atteindre M. Poincaré dans la Note citée plus haut.

Mais il importe essentiellement de remarquer que la démonstration de M. Poincaré s'applique seulement aux points *intérieurs* aux cercles de convergence et non aux points de la circonférence. Pour ceux-ci, le premier des résultats énoncés ne serait pas exact; quant au secoud, il exigerait, semble-t-il, de nouvelles recherches. Sans nous étendre beaucoup sur ce point, considérons une fonction admettant, comme ligne singulière essentielle, un arc de courbe autre qu'un cercle, et convenons d'appeler *valeur de la*

fonction en un point x_0 de l'arc la limite vers laquelle tend sa valeur lorsqu'on s'approche de ce point par un chemin quelconque, non tangent à la ligne singulière, lorsque cette limite existe. Il est clair qu'on ne pourra obtenir cette valeur par la méthode du prolongement analytique, qu'en considérant un élément $\mathfrak{P}(x-a)$, dont le cercle de convergence soit tangent à l'arc au point x_0. On voit que pour obtenir la valeur de la fonction, quand elle existe en tous les points de l'arc, il est nécessaire de considérer une infinité non dénombrable d'éléments [1] $\mathfrak{P}(x-a)$. Cette remarque accroît encore l'importance de la proposition de MM. Poincaré et Volterra, qu'on aurait pu être tenté de considérer comme évidente.

III. — Remarque de Weierstrass sur les séries de fonctions uniformes.

Nous allons maintenant exposer une remarque importante, due à Weierstrass, et relative aux séries dont les termes sont des fonctions uniformes; nous verrons qu'un fait remarquable, qu'il a découvert à l'aide de calculs assez compliqués, est une conséquence immédiate de l'existence de fonctions analytiques non uniformes.

Considérons un élément de fonction analytique, représentant une fonction non uniforme dans une couronne ayant une partie commune avec son cercle de convergence; ce sera, par exemple, la série suivante

$$\mathfrak{P}(x) = \frac{x}{1} + \frac{x^2}{2} + \frac{x^3}{3} + \ldots + \frac{x^n}{n} + \ldots$$

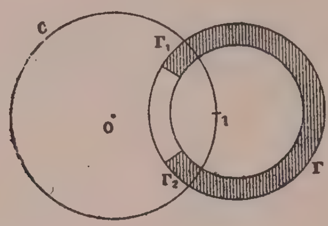
qui représente, comme l'on sait, la fonction $\log\left(\frac{1}{1-x}\right)$, non uniforme dans une couronne entourant le point $x = 1$. D'ailleurs, la série $\mathfrak{P}(x)$ représente la détermination de la fonction qui se réduit à zéro pour $x = 0$.

---

(1) On pourrait croire qu'il suffit de chercher la valeur de la fonction en une infinité dénombrable de points, formant un ensemble partout dense sur l'arc, par exemple aux points dont l'abscisse est rationnelle; mais la valeur de la fonction, telle que nous l'avons définie, n'est pas nécessairement une fonction continue sur l'arc et, par suite, il est absolument indispensable de la calculer en tous les points de cet arc.

Figurons une couronne Γ entourant le point $x = 1$ et supprimons de cette couronne la portion qui n'est pas couverte de hachures (*fig.* 1); il est manifeste que, dans la portion restante, toute branche de la fonction $\log \frac{1}{1-x}$ est uniforme. Considérons l'une de ces branches, par exemple celle dont la partie imaginaire

Fig. 1.



est égale à πi lorsque x est réel; nous aurons défini, dans le domaine D couvert de hachures, *une fonction analytique uniforme bien déterminée.* Désignons cette fonction par $\varphi(x)$. Considérons maintenant la série $\mathfrak{P}(x)$; c'est une série dont les termes sont des fonctions uniformes dans le domaine D et qui d'ailleurs n'y admettent aucun point singulier. Cette série est convergente en deux parties séparées de D, Γ_1 et Γ_2; d'ailleurs, dans chacune d'elles, elle est uniformément convergente. Mais, ce qu'il y a de remarquable, c'est que, dans Γ_1, la somme de la série $\mathfrak{P}(x)$ est égale à $\varphi(x)$, tandis que, dans le domaine Γ_2, la somme de la série $\mathfrak{P}(x)$ est égale à $\varphi(x) - 2i\pi$; c'est une conséquence immédiate des propriétés du logarithme.

En résumé, *nous avons un domaine* D *d'un seul tenant, et, dans ce domaine, une fonction analytique uniforme* $\varphi(x)$; *nous avons d'autre part une série, dont les termes sont des fonctions uniformes n'admettant aucun point singulier dans* D; *cette série converge en deux portions séparées de* D : *dans l'une de ces portions sa somme est* $\varphi(x)$; *dans l'autre, sa somme n'est pas* $\varphi(x)$.

Ce résultat, qui est une conséquence immédiate de l'*existence* de fonctions non uniformes, suffit à mettre nettement en évidence la différence essentielle qu'il y a entre une *expression analytique* et une *fonction analytique;* une même expression analy-

tique, pourvue de sens en deux régions Γ_1 et Γ_2 d'un domaine D d'un seul tenant et égale, en tous les points de Γ_1, à une fonction analytique *uniforme dans* D, n'est pas nécessairement égale à la même fonction dans Γ_2. Admettons maintenant le résultat suivant, qu'on est naturellement conduit à regarder comme vraisemblable et qui est d'ailleurs aisé à démontrer (1) : *étant donné un domaine* D, *limité par une courbe sans points singuliers, et une fonction* $\varphi(x)$, *uniforme dans* D, *on peut trouver une série* $\mathfrak{Q}(x)$ *dont les termes sont des fonctions uniformes dans* D *qui, dans tout ce domaine, converge uniformément* (2) *et a pour somme* $\varphi(x)$.

Considérons alors la série

$$\mathfrak{Q}(x) - \mathfrak{P}(x);$$

cette série est uniformément convergente dans les domaines Γ_1 et Γ_2; sa somme est zéro dans Γ_1 et $2i\pi$ dans Γ_2; d'ailleurs, Γ_1 et Γ_2 font partie d'un domaine D à contour simple *dans lequel les termes de la série sont uniformes et n'ont aucun point singulier*.

Nous voyons ainsi qu'une même série

$$\mathfrak{Q}(x) - \mathfrak{P}(x)$$

peut avoir pour somme, en deux régions d'un domaine D dans lequel ses termes sont uniformes, soit $2i\pi$, soit 0; il est clair d'ailleurs que, si $f(x)$ et $g(x)$ sont deux fonctions uniformes dans D, la série

$$f(x) + \frac{1}{2i\pi}[g(x) - f(s)][\mathfrak{Q}(x) - \mathfrak{P}(x)]$$

a pour somme $f(x)$ dans Γ_1 et $g(x)$ dans Γ_2. Par suite, *étant donnée une série*

$$\Sigma f_i(x)$$

(1) *Voir*, par exemple, PAINLEVÉ, *Sur les lignes singulières des fonctions analytiques* (*Annales de la Faculté des Sciences de Toulouse*, 1888). *Voir* aussi les recherches plus récentes de MM. Hilbert et Painlevé, sur les développements en série de polynomes.

(2) Du moins, la convergence est uniforme dans tout domaine intérieur à D.

dont les termes sont des fonctions uniformes dans un domaine D, *si cette série converge uniformément dans deux portions séparées* Γ_1 *et* Γ_2 *de* D, *on ne peut, si l'on ne fait aucune autre hypothèse, indiquer aucun lien entre les deux fonctions analytiques définies par cette série dans* Γ_1 *et* Γ_2. Ces deux fonctions n'ont aucun rapport nécessaire entre elles.

La méthode par laquelle nous avons obtenu ce résultat est *théoriquement simple*, car elle s'appuie sur des propriétés élémentaires des fonctions analytiques dont les conséquences se présentent naturellement, sans qu'il soit besoin d'aucun artifice de calcul. Mais, d'autre part, la série obtenue est d'une nature compliquée et, si nous pouvons prétendre que notre méthode a l'avantage de ne pas exiger le calcul effectif de ses termes, on peut nous répondre justement que c'est bien heureux, car nous aurions pu être gênés pour faire ce calcul : nous nous contentons de savoir qu'il est possible.

On peut arriver à des résultats analogues à l'aide de séries présentant un caractère tout différent : leur signification théorique est compliquée; le calcul de leurs termes, au contraire, est simple et, si l'on voit moins aisément les raisons théoriques générales du résultat obtenu, on aperçoit, par contre, très nettement, le mécanisme simple de calcul qui fournit ce résultat. Chacun, suivant la nature de son esprit, jugera l'une de ces méthodes préférable à l'autre, c'est-à-dire de nature à mieux lui faire comprendre la *raison* du fait analytique qu'elles font connaître toutes deux. Aussi allons-nous, après avoir indiqué la première, donner un exemple de la seconde, qui est celle de Weierstrass. Nous ne choisirons pas d'ailleurs l'exemple primitif de Weierstrass, dont l'exposé exige d'assez longs calculs, mais un exemple dû à M. Jules Tannery et que Weierstrass a lui-même communiqué à l'Académie des Sciences de Berlin, le 21 février 1881. Il repose sur la remarque bien simple suivante : l'expression

$$\frac{1+x^m}{1-x^m}$$

a, lorsque m augmente indéfiniment, $+1$ comme limite si $|x|<1$ et -1 comme limite si $|x|>1$. Dès lors, si nous désignons par

$$m_1, \quad m_2, \quad \ldots, \quad m_n, \quad \ldots$$

des nombres positifs croissants et si nous considérons une série telle que la somme de ses n premiers termes soit

$$S_n = \frac{1 + x^{m_n}}{1 - x^{m_n}},$$

la somme de cette série sera $+1$ ou -1, suivant que le module de x sera inférieur ou supérieur à *un*.

Il est aisé de calculer le terme général de la série, connaissant la somme de ses n premiers termes ; si, pour fixer les idées, on prend

$$m_n = 2^n,$$

on a

$$u_n = S_n - S_{n-1} = \frac{1 + x^{2^n}}{1 - x^{2^n}} - \frac{1 + x^{2^{n-1}}}{1 - x^{2^{n-1}}} = \frac{1 + x^{2^n} - (1 + x^{2^{n-1}})^2}{1 - x^{2^n}} = \frac{-2x^{2^{n-1}}}{1 - x^{2^n}}$$

et l'on obtient la série

$$\frac{1+x}{1-x} + \frac{2x}{x^2-1} + \frac{2x^2}{x^4-1} + \frac{2x^4}{x^8-1} + \ldots$$

dont la somme est $+1$ et -1 suivant que $|x| < 1$ ou $|x| > 1$. On déduit de ce résultat des conséquences analogues à celles que nous avons développées précédemment ; remarquons cependant que les deux régions dans lesquelles la série a une somme différente sont ici séparées par une ligne (le cercle de rayon *un*) telle qu'un arc quelconque de cette ligne renferme une infinité de points, dont chacun est singulier pour une infinité de termes de la série. A cause de cette circonstance, les résultats qu'on en déduit ne sont pas identiques à ceux que nous avons énoncés tout à l'heure. Mais nous reviendrons sur ce point dans le Chapitre VI ; le Chapitre V est consacré à l'étude de la convergence de certaines séries, étude qui nous est indispensable pour notre but, mais qui ne joue qu'un rôle auxiliaire et que nous avons, par suite, mise à part, afin de ne pas interrompre trop souvent les raisonnements par des calculs accessoires.

CHAPITRE V.

SUR LA CONVERGENCE DE CERTAINES SÉRIES RÉELLES.

I. — Définition des séries étudiées.

Comme nous venons de l'expliquer, les *résultats* seuls de ce Chapitre sont indispensables pour l'intelligence de ce qui va suivre; il n'est pas nécessaire d'en connaître les démonstrations. Le lecteur pour qui ces démonstrations ne présenteraient pas d'intérêt en elles-mêmes peut les omettre sans inconvénient, et lire seulement les résultats que nous résumons à la fin du Chapitre.

Les séries dont nous allons nous occuper sont de la forme suivante :

$$\sum_{n=1}^{\infty} \frac{A_n}{r_n^{m_n}}, \tag{1}$$

les coefficients A_n et les exposants m_n étant des nombres réels positifs; la quantité réelle et *positive* r_n est d'ailleurs définie par l'égalité

$$r_n^2 = (x - a_n)^2 + (y - b_n)^2 + (z - c_n)^2 + \ldots,$$

les variables x, y, z, ... étant en nombre fini quelconque h. De plus, on suppose que la série

$$\Sigma A_n$$

est convergente, et enfin que les exposants m_n sont tous inférieurs à un nombre fixe m. Dès lors, il est clair que si l'on considère, dans l'espace à h dimensions, l'ensemble E des points

$$x = a_n, \qquad y = b_n, \qquad z = c_n, \qquad \ldots$$

et son ensemble dérivé E′, la série (1) est certainement convergente pour tout point x, y, z, ... qui n'appartient pas à E′. Il

existe en effet alors un nombre η tel qu'on ait, pour toute valeur de x,

$$r_n > \eta;$$

on en conclut, m_n étant inférieur à m,

$$r_n^{m_n} > \eta^{m_n} > \eta^m,$$

en supposant ([1]), ce qui est permis, $\eta < 1$. Les termes de la série (1) sont donc respectivement inférieurs aux termes de la série convergente à termes positifs

$$\sum \frac{A_n}{\eta^m}.$$

Le problème que nous nous proposons, c'est l'étude de la convergence de la série pour les points de E'. Nous supposerons d'abord le nombre h des variables égal à *un*, puis égal à *deux*, et l'on verra aisément qu'en se bornant à ce dernier cas on ne restreint qu'en apparence la généralité. Nous supposerons d'abord qu'on a $m_n = 1$, quel que soit n; cette restriction n'est pas non plus essentielle et a seulement pour but de simplifier les notations et le langage.

II. — Séries à une seule variable.

Nous posons donc

$$r_n^2 = (x - a_n)^2,$$

c'est-à-dire

$$r_n = |x - a_n|,$$

et nous considérons tout d'abord la série à termes positifs

$$\sum \frac{A_n}{r_n}. \tag{2}$$

Nous désignons par E l'ensemble des points a_n, par E' l'ensemble dérivé, et nous savons déjà que la série

$$\Sigma A_n \tag{3}$$

([1]) Il est clair que, si m_n n'est pas entier, $r_n^{m_n}$ est supposé essentiellement positif, et par suite a un sens bien précis.

étant supposée convergente, la série (2) est convergente lorsque x n'appartient pas à E' [1]. Il est clair que, pour étudier la convergence de la série (2) lorsque x est compris entre 0 et 1 [2], par exemple, on peut négliger tous les termes correspondant à des valeurs de a_n non comprises dans cet intervalle. Nous supposerons donc que tous les a_n sont compris entre 0 et 1, et nous étudierons la convergence de la série dans cet intervalle. Nous allons donc utiliser pour cela la proposition élémentaire suivante, relative aux séries à termes positifs :

Pour qu'une série soit convergente, il suffit que ses termes soient, à partir d'un certain rang, respectivement inférieurs à ceux d'une série convergente arbitraire, donnée d'avance.

De plus, si les termes de la série étudiée sont fonction d'une ou de plusieurs variables, et si, pour un ensemble de valeurs de ces variables, on peut choisir la même série de comparaison [3], *la série est uniformément convergente pour cet ensemble de valeurs.*

Enfin on peut dire, et c'est presque une tautologie, que, si une série est convergente, on peut trouver une série convergente de comparaison à l'aide de laquelle sa convergence peut être démontrée : il suffit de prendre la série dont chaque terme est le double du terme correspondant de la série donnée. Malgré son extrême banalité, cette remarque ne nous sera pas inutile.

Désignons par

$$(1) \qquad u_1 + u_2 + \ldots + u_n + \ldots$$

la série convergente à laquelle nous allons comparer la série proposée :

$$(2) \qquad \frac{A_1}{r_1} + \frac{A_2}{r_2} + \ldots + \frac{A_n}{r_n} + \ldots.$$

Nous écrirons qu'on a, sauf pour un nombre limité de valeurs

---

(1) Réciproquement, si les a_n sont finis, et si la série (3) est divergente, la série (2) est divergente pour toute valeur de x.

(2) Les limites étant exclues.

(3) On suppose, en outre, que le *rang* à partir duquel les termes de la série étudiée sont inférieurs à ceux de la série de comparaison peut être fixé indépendamment des valeurs particulières des variables.

de n,

$$\frac{A_n}{r_n} < u_n,$$

c'est-à-dire, puisque toutes les lettres désignent des quantités positives,

$$(3) \qquad r_n > \frac{A_n}{u_n}.$$

Ces dernières inégalités ont une interprétation géométrique simple : nous avons posé

$$r_n = |x - a_n|.$$

L'inégalité (3) exprime donc simplement que le point x ne se trouve pas dans l'intervalle, d'étendue $\frac{2A_n}{u_n}$, obtenu en portant de chaque côté du point a_n une longueur égale à $\frac{A_n}{u_n}$. Nous poserons

$$\frac{A_n}{u_n} = v_n,$$

et nous considérerons l'ensemble des intervalles

$$(4) \qquad a_n - v_n, \quad a_n + v_n;$$

l'inégalité (3) devant être vérifiée, quel que soit n, sauf peut-être pour un nombre limité de valeurs, le point x ne doit appartenir qu'à un *nombre fini* [1] des intervalles (4).

Or ces intervalles forment, par leur réunion, un ensemble mesurable, dont la mesure est inférieure ou égale à

$$2\Sigma v_n;$$

nous serons donc *assurés* qu'il existe des points x n'appartenant qu'à un nombre limité de ces intervalles si cette mesure est finie [2], c'est-à-dire si la série

$$(5) \qquad \Sigma v_n$$

---

[1] Il faut évidemment de plus, comme nous l'avons déjà dit, que le point x ne coïncide pas avec un point a_n. Cependant, dans ce cas, nos raisonnements s'appliqueraient à la série dont on aura retranché le terme infini.

[2] Il est clair, en effet, que si chaque point de l'intervalle 0-1 appartient à une infinité d'intervalles, la somme de ces intervalles est infinie.

est convergente. L'étude du cas où cette série ne serait pas convergente est très délicate; nous la laisserons de côté.

Nous supposons donc que la série (5) est convergente; soit v sa somme. Nous allons montrer, dans cette hypothèse, que la série proposée est *uniformément convergente pour un ensemble de valeurs de x dont la mesure peut être supposée aussi voisine qu'on veut de un.* Prenons, en effet, comme série de comparaison,

$$(1') \qquad \Sigma u'_n = \Sigma 2ku_n.$$

Nous avons posé

$$\frac{A_n}{u_n} = v_n$$

et

$$\Sigma v_n = v;$$

nous aurons, en prenant

$$\frac{A^n}{u'_n} = \frac{A_n}{2ku_n} = v'_n,$$

l'égalité fondamentale

$$2\Sigma v'_n = \frac{1}{k}\Sigma v_n = \frac{v}{k}.$$

Or les termes de la série proposée (2) sont certainement tous inférieurs à ceux de la série (1'), si l'on suppose que x n'est situé dans aucun des intervalles

$$(4') \qquad \alpha_n - v'_n, \quad a^n + v'_n,$$

intervalles dont l'étendue totale est

$$2\Sigma v'_n = \frac{v}{k}.$$

L'ensemble des valeurs de x satisfaisant à cette condition a une mesure égale ou supérieure [1] à $1 - \frac{v}{k}$. Or on a pu choisir k de manière que cette expression diffère de 1 aussi peu qu'on veut. Il est clair d'ailleurs que, si l'on considère des valeurs de k de plus en plus grandes, l'ensemble des valeurs de x acquerra sans cesse

---

[1] Elle est supérieure lorsque les intervalles (4') empiètent les uns sur les autres; leur *mesure* est alors inférieure à leur étendue totale.

de nouveaux points ; quelle que soit la valeur fixe de k, la convergence est uniforme pour tous les points de l'ensemble correspondant ; la convergence n'est pas uniforme lorsque l'on considère simultanément toutes les valeurs de k et toutes les valeurs de x qui leur correspondent. Tout ce qu'on peut affirmer, c'est que l'ensemble des valeurs de x pour lesquelles la série n'est pas convergente a pour mesure zéro [1] et que, par suite, l'ensemble des points de convergence est partout dense [2].

Tous ces résultats ont été obtenus en supposant la convergence de la série

$$\Sigma v_n;$$

on avait d'ailleurs posé

$$v_n = \frac{A_n}{u_n}$$

et la série

$$\Sigma u_n$$

était une série convergente quelconque.

Nous avons vu, d'ailleurs, qu'on obtenait certainement la condition nécessaire de convergence de la série proposée, en prenant pour Σu_n une série convergente *arbitraire*, c'est-à-dire en ne faisant pas d'autre hypothèse sur cette série à termes positifs que celle de sa convergence

Nous sommes ainsi conduits à nous poser la question suivante : Étant donnée la série

$$\Sigma A_n$$

(peut-on trouver une série convergente),

$$\Sigma u_n \tag{6}$$

telle que, en posant

$$v_n = \frac{A_n}{u_n}, \tag{7}$$

---

[1] Nous ne sommes d'ailleurs pas assurés que cet ensemble soit mesurable ; en employant le langage expliqué page 48, nous devrions dire que sa mesure est *inférieure ou égale à zéro ;* mais la mesure n'est jamais négative.

[2] Ce dernier point, qui ne suffirait pas à entraîner le résultat d'où nous le tirons, est d'ailleurs évident. Il est clair, en effet, que le raisonnement qui nous a prouvé l'existence des points x dans l'intervalle 0-1 s'appliquerait à tout autre intervalle.

la série

$$(8) \qquad \Sigma v_n$$

soit convergente?

Il est aisé de répondre à cette question; en effet, l'inégalité

$$\sqrt{A_n} = \sqrt{u_n v_n} \leqq \frac{u_n + v_n}{2}$$

nous montre que la convergence des séries (6) et (8) et la relation (7) entraînent la convergence de la série à termes positifs

$$\Sigma \sqrt{A_n};$$

Réciproquement, si cette série est convergente, il suffira de prendre

$$u_n = v_n = \sqrt{A_n},$$

pour que la relation (7) soit vérifiée et que les séries (6) et (8) convergent.

La convergence de la série

$$\Sigma \sqrt{A_n}$$

est donc la condition nécessaire et suffisante pour que nous puissions étudier la série proposée par la méthode simple de la page 66, et cette méthode nous conduit immédiatement au résultat suivant : en posant

$$\Sigma \sqrt{A_n} = v,$$

l'ensemble des points x pour lesquels la série proposée a TOUS *ses termes inférieurs à ceux de la série convergente*

$$\sum \frac{2v}{\varepsilon} \sqrt{A_n}$$

a une mesure supérieure à $1 - \varepsilon$.

Pour tous ces points, la somme de la série est manifestement inférieure à $\frac{2v^2}{\varepsilon}$.

On verrait aisément que, pour la série

$$\sum \frac{A_n}{r_n^{m_n}}, \qquad m_n \leqq m,$$

il suffirait de poser

$$v_n = \frac{A_n}{u_n^m}$$

et de supposer la convergence des deux séries

$$\Sigma v_n$$

et

$$\Sigma u_n,$$

d'où l'on conclurait la convergence de la série

$$\Sigma A_n^{\frac{1}{m+1}}.$$

Cette remarque nous sera utile plus tard.

Séries à deux variables.

Nous allons maintenant étudier le cas où il y a plus d'une variable dans la série ; comme nous l'avons dit, nous supposerons qu'il y en a *deux*, afin de simplifier le langage géométrique.

Nous posons donc maintenant

$$r_n^2 = (x - a_n)^2 + (y - b_n)^2,$$
$$r_n > 0$$

et nous considérons la série à termes positifs

$$\sum \frac{A_n}{r_n};$$

nous écrirons, de même que tout à l'heure,

$$\frac{A_n}{r_n} < u_n$$

et, par suite,

$$(1) \qquad r_n > v_n,$$

en posant

$$v_n = \frac{A_n}{u_n}.$$

La série

$$(2) \qquad \Sigma u_n$$

sera supposée convergente ; mais l'interprétation géométrique des inégalités (1) ne sera plus la même que précédemment. Ces inégalités expriment, en effet, que le point x est extérieur aux divers cercles ayant respectivement pour centres les points a_n, b_n, et pour rayons v_n. Or l'aire de l'un de ces cercles est πv_n^2; la somme de ces aires sera donc finie si l'on suppose convergente la série

$$\Sigma v_n^2. \tag{3}$$

Dès lors, en remplaçant u_n par $k u_n$, on démontrera aisément que la série proposée est uniformément convergente pour tous les points obtenus en supprimant une infinité de cercles dont l'aire totale a pu être supposée aussi petite qu'on veut ; par suite, dans toute région du plan, il y a des points de convergence. D'ailleurs, on voit facilement que la convergence des séries (2) et (3) exige la convergence de la série

$$\Sigma A_n^{\frac{2}{3}}$$

et que, réciproquement, si cette série est convergente, il suffit de prendre

$$u_n = A_n^{\frac{1}{3}}, \qquad v_n = A_n^{\frac{1}{3}},$$

pour pouvoir faire les raisonnements qui précèdent.

Mais nous ne nous étendrons pas sur ces résultats, qui ne nous seront pas utiles. Ils supposent d'ailleurs une étude des ensembles de points à deux dimensions, étude qui, dans ses éléments, n'est pas difficile, mais que nous n'avons pas faite. Nous aurons, au contraire, grand besoin des résultats que nous allons obtenir en étudiant la convergence de la série sur les diverses courbes qu'on peut tracer dans le plan. Nous ferons tout d'abord l'hypothèse que la série

$$\Sigma \sqrt{A_n}$$

est convergente, par suite, nous pourrons supposer que les séries

$$\Sigma u_n$$

et

$$\Sigma v_n$$

sont aussi convergentes.

Dès lors, les cercles C définis par les inégalités (1) seront tels, que *la somme de leurs diamètres pourra être supposée aussi petite qu'on veut* (en changent u_n en ku_n).

Cela posé, considérons une droite quelconque, tracée dans le plan des deux variables réelles x, y. Nous obtiendrons des points de la droite pour lesquels la série est convergente, en supprimant les intervalles intérieurs aux cercles C ; or, toute corde d'un cercle étant plus petite que le diamètre, la somme des segments de la droite qui sont intérieurs aux divers cercles C est inférieure à la somme des diamètres de ces cercles et peut, par suite, être supposée aussi petite qu'on veut. On arrive par suite sans peine, dans l'étude de la convergence de la série sur une droite arbitraire, aux résultats mêmes que nous avions obtenus dans l'étude du cas d'une seule variable; il est inutile de le répéter en détail.

D'ailleurs, ce que nous venons de dire pour une droite s'applique à toute courbe satisfaisant à la condition suivante : la somme des arcs, en chaque point desquels le rayon de courbure est inférieur à ε, tend vers zéro en même temps que ε. En effet, si le rayon de courbure en tout point d'un arc est supérieur à un nombre fixe, on peut affirmer qu'un cercle suffisamment petit intercepte sur cet arc un segment au plus égal au diamètre du cercle [à des infiniment petits près d'ordre supérieur, le diamètre étant regardé comme l'infiniment petit principal ([1])].

Ainsi, en excluant des courbes d'un caractère tout à fait exceptionnel, l'étude de notre série à deux variables sur une courbe quelconque conduit aux mêmes résultats que l'étude de la série à une seule variable, quand on fait la même hypothèse fondamentale : *la série*

$$\Sigma\sqrt{A_n}$$

est convergente.

On aurait pu d'ailleurs déduire ces résultats de ceux que nous avions démontrés pour le cas d'une seule variable, et si notre étude

---

([1]) Nous supposons, bien entendu, que la courbe considérée est telle que les mots *arc*, *rayon de courbure*, ont un sens, et que le rayon de courbure est une fonction continue de l'arc n'admettant point une infinité de maxima et de minima. On pourrait s'affranchir d'une partie de ces restrictions, mais cela n'a pour nous aucun intérêt.

se bornait là, elle ne fournirait pas une véritable généralisation de ces premiers résultats. Nous allons obtenir, au contraire, des conséquences vraiment nouvelles et importantes pour la suite, en considérant, non plus une courbe isolée, mais une famille de courbes.

Courbes de convergence uniforme.

Proposons-nous d'abord de rechercher comment se comporte la série proposée,

$$\sum \frac{A_n}{r_n},$$

sur un ensemble de droites parallèles D ; nous supposons toujours, bien entendu, que la série $\Sigma\sqrt{A_n}$ est convergente. Soit D′ une droite perpendiculaire aux droites D; chaque droite D est déterminée par le point où elle rencontre D′. Si nous considérons un des cercles C, les droites D qui le rencontrent déterminent sur D′ un segment dont la longueur est égale au diamètre du cercle considéré : c'est sa projection orthogonale sur D′. A l'ensemble des cercles C correspond ainsi un ensemble de segments, dont la somme peut être rendue aussi petite qu'on veut, puisqu'elle est égale à la somme des diamètres des cercles.

Mais, d'après la manière même dont ces segments ont été construits, il est clair que toute droite D, déterminée par un point de D′ n'appartenant à aucun de ces segments, ne rencontre aucun des cercles C. Par conséquent, en choisissant convenablement la constante k, dont nous avons déjà fait usage, nous pourrons affirmer qu'il existe sur un segment quelconque de D′ un ensemble E, dont la mesure est aussi voisine qu'on veut de la longueur totale de ce segment et qui, de plus, est tel que, sur toutes les droites D déterminées par les points de E, *la série est uniformément convergente*. Car, sur toutes ces droites, on a pour toute valeur de n

$$\frac{A_n}{r_n} < k\sqrt{A_n},$$

k étant un nombre fixe et la série

$$\Sigma\sqrt{A_n}$$

étant convergente. D'ailleurs, si l'on pose

$$\Sigma \sqrt{A_n} = v$$

et si l'on considère un segment de D' dont la longueur est l, on pourra affirmer que la mesure de l'ensemble E est supérieure à

$$l - \frac{2v}{k}.$$

Comme nous l'avons dit plus haut dans des circonstances analogues, on peut affirmer que la convergence a lieu pour un ensemble de mesure l, mais elle n'est pas uniforme pour cet ensemble, bien qu'elle le soit sur chaque droite D.

L'importance du résultat précédent dans les applications est due à ce fait que la série convergeant uniformément sur certaines droites D, sa somme est une fonction continue de la position du point x, y lorsque ce point décrit l'une de ces droites. On ne pouvait évidemment rien affirmer de pareil, pour une série convergeant uniformément pour les points d'un ensemble non continu.

Il est d'ailleurs aisé de généraliser le résultat précédent, en remplaçant la famille de droites parallèles D par une famille de courbes Γ satisfaisant à des conditions très générales. Voici ces conditions restrictives dont plusieurs ne sont pas indispensables ou font double emploi, mais qui ne nous gêneront pas pour les applications. Nous supposons qu'on peut associer à la famille Γ une famille Γ' telle que, dans une région A du plan, ces deux familles de courbes déterminent un réseau de parallélogrammes infiniment petits sans *point singulier*. Nos conclusions s'appliquent alors à la région A; elles s'appliqueraient à tout le plan si tous les points a' étaient intérieurs à la région A.

Voici ce que nous entendons par un réseau sans point singulier; on a, *dans la région* A,

$$x = \varphi(u, v),$$
$$y = \psi(u, v);$$

à tout système de valeur de u, v correspond un seul système de valeurs de x, y et réciproquement; les courbes Γ et Γ' s'obtiennent respectivement en prenant $u =$ const., $v =$ const., et en

posant

$$dx^2 + dy^2 = \mathrm{E}\, du^2 + 2\mathrm{F}\, du\, dv + \mathrm{G}\, dv^2,$$

les quantités essentiellement positives

$$\mathrm{E}, \quad \mathrm{G}, \quad \mathrm{EG} - \mathrm{F}^2$$

sont, dans toute la région A, comprises entre deux nombres positifs fixes.

On voit aisément alors que l'une quelconque des courbes Γ' peut jouer le même rôle que la droite D' dans la démonstration précédente : les courbes Γ qui rencontrent un cercle C interceptent sur Γ' un segment dont le rapport au diamètre du cercle est compris entre deux nombre fixes et l'on démontre ainsi aisément que les courbes Γ ont les mêmes propriétés essentielles que les droites D, au point de vue de la convergence de la série proposée.

Sans examiner tous les cas où les courbes Γ ne satisfont pas aux conditions exigées, ce qui a lieu, par exemple, si, dans l'aire A considérée, elles ont une enveloppe, nous allons étudier un cas particulier important, celui où les courbes Γ sont des droites passant par un point fixe. Il est clair d'ailleurs que, si ce point fixe n'était pas un point a', nous rentrerions dans un cas déjà étudié ; nous supposons donc que nous avons des droites Δ passant par un même point x, lequel est un point limite de l'ensemble des a. La première question qui se pose, dans l'ordre d'idées où nous sommes restés jusqu'ici, est la suivante : Y a-t-il des droites Δ sur lesquelles la série est convergente? Nous remplacerons cette question, par la suivante, plus restrictive : Y a-t-il des droites Δ qui ne rencontrent aucun des cercles C? Nous supposons d'ailleurs le système des cercles C défini de la manière la plus générale ; *le rayon du cercle* C_n *est* v_n *et la série*

$$\sum \frac{A_n}{v_n}$$

est convergente ; ce sont là, pour le moment, nos seules hypothèses.

Il est clair d'abord que le point x doit être extérieur à tous les cercles C; sinon, toute droite Δ rencontrerait l'un au moins de ces cercles; on a donc certainement, pour toute valeur de n,

$$r_n > v_n.$$

Menons du point x les deux tangentes au cercle C_n ; l'angle de ces deux tangentes est visiblement

$$2 \arc \sin \frac{v_n}{r_n},$$

le symbole *arc sin* désignant un arc compris entre o et $\frac{\pi}{2}$; d'ailleurs, il est clair qu'une droite Δ rencontre ou ne rencontre pas le cercle C_n suivant qu'elle est inférieure ou extérieure à cet angle.

On conclut aisément de là que, pour qu'il existe des droites Δ ne rencontrant aucun des cercles C_n, il *suffit* (¹) qu'on ait

$$\sum_1^\infty 2 \arc \sin \frac{v_n}{r_n} < 2\pi. \tag{1}$$

Il est donc tout d'abord nécessaire que la série

$$\sum \frac{v_n}{r_n}$$

soit convergente et il est manifeste que cette condition suffit, car, si elle est remplie, on pourra toujours vérifier l'inégalité (1) en remplaçant v_n par kv_n et choisissant convenablement la constante k. On pourra de même démontrer qu'il existe des droites Δ dans un angle aussi petit qu'on veut, ayant son sommet au point x. Tout cela est une conséquence de la convergence de la série

$$\sum \frac{v_n}{r_n};$$

nous allons indiquer un cas très étendu dans lequel on peut trouver une infinité de points x tels que cette série soit conver-

---

(¹) L'inégalité que nous écrivons exprime qu'il existe des *demi-droites*, issues du point x et ne rencontrant aucun des cercles C_n ; mais leur prolongement au delà de x peut en rencontrer; pour qu'il n'y ait pas de *droites* passant par x et ne rencontrant aucun de ces cercles, il faut écrire

$$4 \Sigma \arc \sin \frac{v_n}{r_n} < 2\pi,$$

mais ce facteur numérique est sans importance.

gente. Il suffit que la série

$$\Sigma A_n^{\frac{2}{5}}$$

soit convergente; on pourra en effet, alors, trouver une infinité de points x tels qu'on ait

$$r_n > k A_n^{\frac{1}{5}},$$

k étant une constante convenablement choisie; en prenant d'ailleurs

$$v_n = k' A_n^{\frac{3}{5}},$$

les deux séries

$$\sum \frac{A_n}{v_n}$$

et

$$\sum \frac{v_n}{r_n}$$

seront convergentes, ce qui est le résultat désiré.

Par conséquent, *lorsque la série*

$$\Sigma A_n^{\frac{2}{5}}$$

est convergente, on peut trouver, dans toute région, une infinité de points x, tels que par chacun de ces points passe une infinité de droites sur chacune desquelles la série proposée est convergente. L'ensemble de ces droites a d'ailleurs une mesure angulaire aussi voisine qu'on veut de π, si on les considère comme prolongées dans les deux sens, ou de 2π, si on les regarde comme des demi-droites.

Comme nous l'avons dit plus haut, ces résultats s'étendent sans peine à une série de la forme

$$\text{(1)} \qquad \sum \frac{A_n}{r_n^{m_n}}, \qquad m_n < m;$$

seulement, au lieu de supposer la convergence de l'une des séries

$$\Sigma A_n^{\frac{2}{3}}, \quad \Sigma A_n^{\frac{1}{2}}, \quad \Sigma A_n^{\frac{2}{5}},$$

il faut, suivant le but que l'on désire atteindre, supposer la convergence de la série

$$\Sigma A_n^{p},$$

où p est une fonction de m, aisée à déterminer ; d'ailleurs, p tend vers zéro lorsque m augmente indéfiniment. Par exemple, la convergence de la série

$$\Sigma A_n^{\frac{1}{m+1}}$$

entraîne, pour la série (1), les mêmes conséquences que la convergence de la série $\Sigma A_n^{\frac{1}{2}}$ entraînait pour la série

$$\sum \frac{A_n}{r_n}.$$

Une application importante de ces remarques est la suivante : *Supposons que la série*

$$\Sigma A_n^p$$

soit convergente, quel que soit le nombre positif fixe p ; c'est ce qui a lieu certainement si l'on suppose convergente la série

$$\sum \frac{1}{\left[\log \frac{1}{A_n}\right]^\mu},$$

qui est à termes positifs à partir d'un certain rang, puisque A_n tend vers zéro lorsque n augmente indéfiniment ; μ est un nombre positif quelconque.

Considérons dans cette dernière hypothèse, avec la série

$$\sum \frac{A_n}{r_n^{m_n}}, \qquad m_n \leq m,$$

toutes les séries qu'on en déduit en prenant les dérivées de chaque terme un nombre quelconque de fois, soit par rapport à x, soit par rapport à y. Ces dérivées s'expriment d'ailleurs linéairement au moyen des séries obtenues en dérivant chaque terme par rapport à r_n. Nous pourrons affirmer que, si la série

$$\sum \frac{1}{\left[\log \frac{1}{A_n}\right]^\mu}$$

est convergente, chacune de ces dérivées a toutes les propriétés de convergence que nous avons énoncées plus haut. On pourrait même démontrer des propositions un peu plus précises, par

exemple celle-ci : on peut trouver des courbes telles que, par chacun de leurs points, il passe une infinité de courbes sur lesquelles la série est convergente, etc. Mais nous préférons insister sur un fait plus important; nous venons de dire que chacune des séries dérivées est convergente sur une infinité de courbes; cela ne permet pas de conclure qu'il y a des courbes sur lesquelles toutes les séries dérivées sont convergentes. On arrivera à ce dernier résultat en prenant pour r_n les termes de la série convergente

$$\sum \frac{1}{\left[\log \frac{1}{A_n}\right]^\mu},$$

multipliés au besoin par une constante k, et en raisonnant comme précédemment. On trouvera ainsi, dans les mêmes conditions que plus haut (p. 73), des courbes sur lesquelles *toutes* les séries dérivées sont convergentes; d'ailleurs, sur chaque courbe la convergence est uniforme, d'où l'on conclut aisément que la dérivée de la fonction par rapport à l'arc de la courbe peut se calculer en appliquant simplement la règle de dérivation des séries.

Il suffit pour cela de se rappeler qu'une série uniformément convergente peut être intégrée terme à terme.

Nous allons, en terminant, résumer ceux des résultats acquis dans ce Chapitre dont nous aurons besoin plus loin.

Étant donnée une série de la forme

$$\varphi(x, y) = \sum \frac{A_n}{[(x-a_n)^2+(y-b_n)^2]^{m_n}},$$

on suppose que les A_n *et les* m_n *sont des nombres positifs; les* m_n *sont inférieurs à un nombre fixe et les* A_n *sont tels que la série*

$$\sum \frac{1}{\left[\log \frac{1}{A_n}\right]^\mu},$$

dans laquelle μ *désigne un nombre positif, est convergente* [1].

---

[1] On suppose que les A_n tendent vers zéro; cette dernière série pourrait aussi être convergente si A_n augmentait indéfiniment avec n.

En même temps que la série $\varphi(x, y)$, *on considère les séries obtenues en prenant les dérivées partielles d'ordre quelconque de tous ses termes.*

Cela étant, on peut trouver dans toute région du plan une infinité non dénombrable de courbes C *sur lesquelles toutes ces séries à termes positifs sont uniformément convergentes. Si l'on considère toutes les droites* D *parallèles à une direction donnée et un segment de droite* σ *de longueur* l *non parallèle à cette direction, on peut trouver sur ce segment un ensemble dont la mesure diffère aussi peu qu'on veut de* l *et tel que toute droite* D *passant par un point de cet ensemble ait la propriété fondamentale des courbes* C.

Il en est de même lorsque l'on considère les droites D *passant par un point* O *et qu'on remplace le segment* σ *par un arc de cercle de centre* O, *si l'on suppose que le point* O *ne coïncide avec aucun des points* $(x = a_n, y = b_n)$, *ni avec aucun point de l'ensemble dérivé de l'ensemble formé par ces points. On peut cependant, dans toute région du plan, même si tous les points de cette région appartiennent à l'ensemble dérivé dont nous venons de parler, trouver une infinité non dénombrable de points pouvant être pris pour le point* O. *On peut aussi trouver une infinité non dénombrable de tels points sur toute courbe* C.

CHAPITRE VI.

LA NOTION DE FONCTION D'UNE VARIABLE COMPLEXE.

C'est à Cauchy que revient la gloire d'avoir fondé la théorie des fonctions d'une variable complexe, en faisant connaître notamment la relation qu'il y a entre le rayon de convergence de la série de Taylor et la situation des singularités de la fonction ([1]). D'ailleurs, là comme dans plusieurs autres parties de la Science, Cauchy a laissé beaucoup à faire à ses successeurs; mais on a pu dire que chaque nouvelle découverte dans la théorie des fonctions ajoute, s'il est possible, quelque chose à l'idée que nous avons de son génie. Nous ne pouvons faire ici l'historique du développement immense qu'a pris dans la seconde moitié de ce siècle la Science nouvelle créée par Cauchy; contentons-nous de citer les noms célèbres de Puiseux, Riemann, Weierstrass. L'exposition des idées de Riemann trouverait plutôt sa place dans des leçons sur les fonctions algébriques; nous prendrons pour guide, dans ce Chapitre, les célèbres Mémoires réunis par Weierstrass sous le

([1]) Dans une Notice des plus intéressantes sur Weierstrass (*Acta mathematica*, t. XXI), M. Mittag-Leffler paraît affirmer que le développement des idées de Weierstrass est indépendant des travaux de Cauchy. Il nous semble qu'il importe peu que Weierstrass ait ou n'ait pas connu tel Mémoire particulier de Cauchy; mais il n'est pas douteux qu'à l'époque où Weierstrass a pu commencer à étudier les Mathématiques, la notoriété de Cauchy était trop grande, ses travaux déjà publiés trop nombreux et trop répandus, pour qu'il fût possible de ne pas subir l'influence de sa pensée. De même, actuellement, il serait impossible de *penser* sur la théorie des fonctions en faisant abstraction complète de l'influence de Weierstrass, ne lirait-on pas une seule ligne écrite par lui. La part de Weierstrass dans le développement de la théorie est assez grande et assez belle pour qu'il soit inutile de chercher à l'augmenter en enlevant à Cauchy le mérite d'avoir été là, comme sur bien d'autres points, un initiateur et un précurseur.

titre commun : *Abhandlungen aus der Functionenlehre*, en tenant compte de divers travaux ultérieurs et notamment de la Thèse (¹) de M. Painlevé.

Fonctions analytiques et expressions analytiques.

L'une des idées fondamentales que nous devons à Weierstrass, c'est la distinction précise entre les deux notions de *fonction analytique* et d'*expression analytique*. Nous savons déjà ce qu'on doit entendre, d'après Weierstrass, par *fonction analytique :* c'est l'ensemble de tous les *éléments* $\mathfrak{P}(x-a)$ qui se déduisent les uns des autres par la méthode du *prolongement*. Quant aux mots *expression analytique*, ils désignent, dans leur sens le plus général, toute expression dont on sait, au moyen d'opérations analytiques connues (addition, multiplication, division, intégration définie ou indéfinie, etc.), calculer la valeur quand on connaît la valeur de la variable (²). Le nombre des opérations peut d'ailleurs être fini ou infini; un exemple de ce dernier cas nous est fourni par la considération des séries ou des produits infinis. On sait d'ailleurs que, sous des conditions très larges, une expression analytique représente dans un certain domaine une fonction analytique (³); d'autre part, toute fonction analytique peut être représentée par une expression analytique (par exemple par la série de Taylor, ou par l'intégrale de Cauchy), au moins dans une partie de son domaine naturel d'existence.

Néanmoins, ces deux notions se distinguent nettement, car la notion de fonction analytique est plus précise et la notion d'expression analytique plus large et en quelque mesure plus naturelle. L'importance de la notion de fonction analytique est due surtout à ce fait que deux fonctions analytiques ne peuvent coïncider le long d'un arc de courbe fini sans coïncider dans tout leur domaine

---

(¹) *Sur les lignes singulières des fonctions analytiques* (*Annales de la Faculté des Sciences de Toulouse*, 1888).

(²) M. Mittag-Leffler a donné parfois, aux mots *expression analytique*, un sens plus étendu que nous n'avons pas cru devoir adopter ici.

(³) Par exemple, toute série dont les termes sont des fonctions analytiques représente une fonction analytique dans un domaine d'un seul tenant, si elle converge uniformément dans ce domaine.

d'existence [1]; au contraire, nous avons vu qu'une expression analytique peut avoir des régions diverses de convergence et être nulle dans l'une seulement de ces régions.

On peut se borner à considérer des fonctions analytiques et les définir uniquement par des séries de Taylor; c'est le point de vue auquel se place systématiquement M. Méray; mais on ne peut se dissimuler que ce procédé qui a, il faut le reconnaître, une très grande beauté théorique, présente des inconvénients considérables. D'abord, pour définir une fonction analytique dans tout son domaine d'existence, il faut en général une infinité d'éléments $\mathfrak{P}(x-a)$ et chacune de ces séries $\mathfrak{P}(x-a)$ diverge en des points où la fonction existe. Nous avons vu, d'après M. Poincaré, que, si l'on exclut les points qui forment la frontière du domaine naturel d'existence, on peut se borner à considérer une infinité *dénombrable* d'éléments; mais, dans certains cas, la considération de la valeur de la fonction en ces points peut rendre des services et il est alors nécessaire de considérer une infinité *non dénombrable* d'éléments.

Enfin, et ce dernier inconvénient est peut-être le plus grave, le développement de Taylor $\mathfrak{P}(x-a)$ ne met nullement en évidence les propriétés essentielles de la fonction qu'on étudie, précisément parce qu'il s'applique à toutes les fonctions. La connaissance de son domaine de convergence nous fait connaître seulement la distance au point a du point singulier le plus voisin, et il est extrêmement difficile d'obtenir d'autres renseignements sur la fonction par l'étude de son développement de Taylor; c'est seulement dans des cas extrêmement particuliers [2] qu'on peut tirer

[1] Ce théorème, aisé à démontrer lorsque l'arc de courbe est *analytique*, devient beaucoup plus délicat à seulement énoncer lorsque l'arc de courbe est quelconque et que les fonctions sont définies d'un côté seulement de cet arc. Il nous semble qu'il se transforme alors en la proposition suivante due à M. Painlevé (*Thèse*, p. 28-29) : Étant donnée une fonction analytique définie dans une aire A, limitée par une courbe C, si cette fonction est nulle sur un arc de C, elle est nulle dans A. Dans cet énoncé la courbe C est un ensemble de points tels que le plan se trouve divisé en deux régions A et A′ telles que l'on ne peut aller par un chemin continu, d'un point de A à un point de A′, sans que ce chemin renferme au moins un point de C; un arc de C se définit d'une manière analogue.

[2] Citons cependant un théorème très remarquable dû à M. Hadamard (*Comptes rendus des séances de l'Académie des Sciences*, t. CXXIV, p. 1503)

quelque chose de cette étude, grâce aux brillantes et profondes recherches de MM. Hadamard et Fabry.

Aussi est-il naturel de se poser la question suivante : Étant donnée une fonction analytique, peut-on trouver une expression analytique *qui converge en tout point du domaine naturel d'existence de la fonction* (1), et qui représente cette fonction? Il est d'abord clair que si l'expression analytique est *uniforme*, c'est-à-dire n'a qu'une valeur pour chaque valeur de la variable, elle ne peut représenter qu'une fonction uniforme; il est donc naturel de se borner, au moins dans une première étude, aux fonctions analytiques uniformes (2). D'ailleurs, un remarquable théorème de M. Poincaré (3) nous apprend que, si une variable complexe y est une fonction non uniforme de nature quelconque d'une variable complexe x, on peut exprimer x et y en fonction uniforme d'une même variable complexe t. Les fonctions non uniformes sont ainsi ramenées, à un certain point de vue, aux fonctions uniformes.

Nous énoncerons donc ainsi, en la restreignant, la question que nous posions tout à l'heure : étant donnée une fonction analytique uniforme ayant un domaine naturel d'existence D, absolument quelconque, peut-on trouver une expression analytique uniforme, qui représente cette fonction en tout point de D? Sous sa forme la plus générale, la question a été résolue pour la première fois par M. Runge (*Acta mathematica*, t. VI). Plus récemment,

---

qui fait connaître une *relation* entre les singularités de développements de Taylor tout à fait généraux et liés simplement les uns aux autres.

(1) *Cf.* Weierstrass (*loc. cit.*), *Werke*, t. II, p. 83 ... *ob es möglich sei, arithmetische, aus der Veränderlichen x und aus unbestimmten Constanten zusammengesetzte Ausdrücke zu bilden, welche sämmtliche Functionen einer bestimmten Klasse — und nur diese — darstellen.* On voit que la question posée par Weierstrass est à la fois moins générale et plus précise que la nôtre. Nous reviendrons sur ce point à la fin du Chapitre.

(2) A ma connaissance, le seul résultat général sur la représentation analytique des fonctions analytiques non uniformes est dû à M. Picard (*Mémoire sur les fonctions entières* et *Traité d'Analyse*, t. II). M. Picard suppose qu'on est dans le voisinage d'un point *isolé* de non-uniformité. On pourrait évidemment étendre beaucoup ce résultat par des méthodes analogues à celles de M. Mittag-Leffler que nous citerons bientôt; mais cette extension facile serait dépourvue d'intérêt, à moins qu'on n'obtienne un résultat complètement général, ce qui paraît exiger des considérations assez délicates et en tout cas très longues.

(3) Poincaré, *Sur un théorème de la théorie générale des fonctions* (*Bulletin de la Société mathématique de France*, t. XI).

M. Painlevé en a repris l'étude par des méthodes nouvelles et a publié des résultats importants et très généraux [1] (*Comptes rendus*, janvier et février 1898, *passim*).

De ces diverses recherches nous retiendrons notamment le résultat suivant : *On peut trouver, d'une infinité de manières, une série de fractions rationnelles qui converge absolument et uniformément dans tout domaine* D′ *intérieur à* D *et qui représente la fonction dans tout le domaine* D. La question que nous avions posée se trouve ainsi résolue par l'affirmative; mais les théorèmes de M. Runge et de M. Painlevé ne fournissent pas seulement la réponse à cette importante question; ils soulèvent d'autres questions non moins intéressantes. Du moment, en effet, où l'on est assuré qu'il existe, pour chaque fonction uniforme, une infinité d'expressions analytiques uniformes propres à la représenter dans tout son domaine naturel d'existence, il est naturel de chercher à choisir parmi ces expressions. Le point de vue auquel on devra se placer ressort de la citation de Weierstrass que nous avons faite tout à l'heure : l'expression qu'on choisira doit mettre en évidence le mieux possible les propriétés caractéristiques de la fonction, et en particulier ses singularités. Lorsque l'ensemble des singularités n'est pas dénombrable, on peut penser que la méthode de M. Runge ou celle de M. Painlevé les mettent en évidence autant qu'il est possible de le faire. On ne peut, en effet, songer à faire figurer explicitement dans une expression analytique une infinité non dénombrable de singularités; il semble, dès lors, que tout ce qu'on peut exiger, c'est que *l'ensemble* $\mathcal{E}$ *des singularités de la fonction* coïncide avec l'ensemble obtenu en adjoignant à l'ensemble E *des singularités de l'expression analytique*, son ensemble dérivé E′. Mais il y a une infinité d'ensembles E ayant cette propriété (que E + E′ coïncide avec $\mathcal{E}$); dans bien des cas, l'un de ces ensembles E se distingue nettement de tous les autres et doit être regardé comme formé des véritables singularités de la fonction, les points de E′ n'étant pas singuliers en eux-mêmes, si l'on peut ainsi dire, mais seulement comme limites de points singuliers.

---

(1) *Voir* la Note de M. Painlevé dans les *Leçons sur les fonctions de variable réelle et les développements en séries de polynomes* (Note de la deuxième édition).

Le théorème de M. Mittag-Leffler.

Le problème qui nous occupe a reçu depuis longtemps de M. Mittag-Leffler une solution moins générale que celle de M. Runge ou de M. Painlevé, lesquelles s'appliquent à *toutes* les fonctions uniformes, mais bien plus précise dans les cas, d'ailleurs très étendus, où l'on peut l'utiliser. Indiquons *l'un de ces cas :* soient E l'ensemble des points singuliers d'une fonction uniforme, E_1 l'ensemble dérivé de E, E_2 l'ensemble dérivé de $E_1, \ldots, E_{n+1}$ l'ensemble dérivé de $E_n, \ldots$. S'il existe un nombre p tel qu'on ait

$$E_p = 0,$$

c'est-à-dire tel que l'ensemble E_p ne renferme aucun point, la méthode de M. Mittag-Leffler (¹) s'applique et l'on peut trouver une expression analytique qui représente la fonction en tous les points non singuliers *et dans laquelle chaque point singulier est mis en évidence, ainsi que la nature de la singularité de la fonction en ce point.*

Nous ne pouvons étudier en détail la méthode générale de M. Mittag-Leffler, mais, en prenant pour exemple le cas particulier traité par Weierstrass et qui fut l'origine des recherches de M. Mittag-Leffler, nous allons chercher à préciser dans quelle mesure on peut dire que la fonction uniforme est déterminée par la connaissance de ses singularités.

Considérons, avec Weierstrass, la dérivée logarithmique d'une fonction entière; c'est une fonction uniforme, n'admettant à distance finie d'autres singularités que des pôles simples; chaque pôle a pour résidu $+1$ si les zéros de la fonction entière considérée sont simples.

Donnons-nous une infinité de points

$$a_1, \quad a_2, \quad \ldots, \quad a_n, \quad \ldots,$$

et proposons-nous de former une fonction uniforme admettant ces points comme pôles simples, le résidu de chaque pôle étant $+1$.

(¹) *Sur la représentation analytique des fonctions monogènes uniformes d'une variable indépendante* (*Acta mathematica*, t. IV).

Nous pouvons supposer que les a_n sont rangés de manière que leurs modules n'aillent pas en décroissant; il sera d'abord nécessaire de supposer que $|a_n|$ augmente indéfiniment avec n. Cette condition étant remplie, si la série

$$\sum \frac{1}{|a_n|}$$

est convergente, on peut prendre pour la fonction cherchée l'expression analytique

$$(1) \qquad \sum \frac{1}{z - a_n}.$$

Mais ce n'est évidemment pas là l'expression analytique la plus générale répondant à la question; il est clair que, si l'on désigne par $G(z)$ une fonction entière arbitraire, on peut prendre aussi bien comme solution l'expression suivante :

$$(2) \qquad G(z) + \sum \frac{1}{z - a_n}.$$

Ainsi, le problème que nous nous étions posé est indéterminé; mais il est aisé de le rendre déterminé en complétant son énoncé : il suffit d'exiger que l'expression analytique cherchée soit la dérivée logarithmique d'une fonction entière *de genre zéro*. Dès lors, l'expression (1) fournit la solution unique de ce nouveau problème (à une constante additive près).

Or, en disant que la fonction cherchée était la dérivée logarithmique d'une fonction de genre zéro, on a fixé en quelque mesure la manière dont elle se comporte à l'infini [1]; cela a suffi pour la déterminer, grâce à cette hypothèse préalable, *la convergence de la série*

$$\sum \frac{1}{|a_n|}.$$

Si l'on n'avait d'abord supposé cette convergence, on n'aurait pu arriver aussi simplement à rendre unique la solution du problème. Pour nous rendre compte nettement de ce fait, examinons

---

[1] On peut rapprocher les considérations du texte d'une Note que j'ai publiée *sur l'interpolation* (*Comptes rendus*, mars 1897).

le cas où la série

$$\sum \frac{1}{|a_n|^{p+1}}$$

est convergente, la série

$$\sum \frac{1}{|a_n|^p}$$

étant divergente; une solution particulière du problème est alors fournie par la série

$$(3) \qquad \sum \left[\frac{1}{z - a_n} + \frac{1}{a_n} + \frac{z}{a_n^2} + \ldots + \frac{z^{p-1}}{a_n^p} \right]$$

et la solution la plus générale, par l'expression analytique

$$(4) \qquad G(z) + \sum \left[\frac{1}{z - a_n} + \frac{1}{a_n} + \ldots + \frac{z^{p-1}}{a_n^p} \right].$$

Est-il possible ici, au moyen de conditions supplémentaires, relatives à la manière dont la fonction se comporte à l'infini, de distinguer parmi l'infinité des solutions que renferme la formule (4). Il semble d'abord que la solution (3) se sépare nettement des autres; il suffit de changer z en $z - z_0$ pour constater que cette distinction est en partie artificielle. Mais, si nous assujettissons la fonction cherchée à être la dérivée logarithmique d'une fonction entière de genre p, nous pourrons affirmer que $G(z)$ est un polynome de degré au plus égal à p et l'expression (4) ne renfermera plus que $p + 1$ constantes arbitraires (dont l'une peut être négligée). Il suffira donc d'avoir, sur la manière dont se comporte la fonction à l'infini, des renseignements qui permettent d'écrire p relations entre ces constantes inconnues pour que la fonction soit entièrement déterminée, à une constante additive près. On voit que cette solution est moins satisfaisante que dans le cas où le genre est égal à zéro. Mais on peut dire tout au moins que, même si l'on ne cherche pas à préciser complètement la solution, la seule hypothèse que le genre est p suffit à diminuer considérablement l'indétermination, puisqu'il ne subsiste plus que p constantes, au lieu d'une infinité.

Il n'en est plus du tout de même lorsqu'il n'existe pas de nombre p tel que la série

$$\sum \frac{1}{|a_n|^p}$$

soit convergente; on sait que, dans ce cas, on est obligé, pour avoir une solution du problème, d'introduire des polynomes dont le degré dépasse toute limite; on peut, *par exemple*, considérer l'expression

$$\sum\left(\frac{1}{z-a_n}+\frac{1}{a_n}+\frac{z}{a_n^2}+\ldots+\frac{z^{n-1}}{a_n^n}\right);$$

mais il est aisé de voir qu'il n'y a plus aucune manière qui ne soit artificielle, pour distinguer cette expression de celles qu'on obtient en lui ajoutant une fonction entière arbitraire. S'il est possible de *déterminer* la solution au moyen de conditions supplémentaires, c'est bien probablement au moyen de considérations compliquées et très délicates.

Les difficultés que nous venons de signaler croîtraient beaucoup si l'on étudiait, au point de vue de la *détermination* de la fonction au moyen de ses singularités, le cas le plus général traité par M. Mittag-Leffler, dans le Mémoire que nous avons cité. Néanmoins, si l'on peut reprocher à la solution de M. Mittag-Leffler le défaut de renfermer un certain arbitraire, et cet arbitraire est peut-être impossible ou extrêmement difficile à ne pas introduire, on peut dire que les éléments *essentiels* de la représentation analytique sont déterminés par la connaissance des singularités de la fonction et ce fait suffit pour distinguer nettement cette représentation de celle de M. Runge dans laquelle les pôles des fonctions rationnelles introduites sont à peu près complètement arbitraires (¹). Mais cette dernière a, par contre, l'avantage d'être absolument générale.

Les représentations analytiques connues.

Pour résumer les résultats acquis sur le problème de la représentation analytique des fonctions uniformes, nous pouvons dire

(¹) La même objection peut être faite aux divers développements de M. Painlevé, que nous avons déjà cités. Le développement sous forme de produit infini paraît être celui où les propriétés de la fonction sont le mieux mises en évidence. D'ailleurs, M. Painlevé se proposait un but tout différent du nôtre: au lieu de rechercher un développement unique, il a cherché, semble-t-il, à introduire le plus d'arbitraire possible dans sa solution, ce qui peut être très commode dans certaines applications.

que nous en connaissons deux solutions complètes; l'une est fournie par le théorème de Taylor, l'autre par le théorème de M. Runge ([1]). Ces deux solutions ont une très grande importance à cause de leur généralité; mais chacune d'elles a de graves inconvénients dont les principaux sont, pour la série de Taylor, de diverger en des régions où la fonction existe; et, pour la représentation de M. Runge et celles de M. Painlevé, d'être possibles d'une infinité de manières.

Indépendamment de ces deux solutions générales, on connaît des représentations analytiques qui ne s'appliquent qu'à des classes particulières de fonctions, mais qui, dans les cas où l'on peut les employer, sont très utiles, parce qu'elles mettent en évidence les propriétés de la fonction. Parmi ces représentations particulières, la plus importante est celle de M. Mittag-Leffler dont nous avons déjà parlé.

Il est donc naturel de se proposer le problème suivant : rechercher pour les fonctions uniformes une représentation analytique qui mette en évidence leurs propriétés essentielles et qui, pour une fonction donnée, soit unique; de telle sorte que les expressions analytiques, égales à zéro dans toute une région du plan, sans être identiquement nulles, soient exclues. On exige de plus que l'expression analytique converge dans une région aussi étendue que possible.

Il semble qu'on ne peut guère espérer résoudre d'un seul coup ce problème pour toutes les fonctions; il est donc nécessaire de le restreindre en faisant des hypothèses sur les singularités des fonctions étudiées. A cause de la diversité des hypothèses possibles, il y a là un champ considérable de recherches; c'est à force d'étudier des cas particuliers, parmi ceux qui paraissent les plus intéressants, qu'on arrivera peut-être à avoir quelques idées sur la solution du problème général.

On peut poser la question sous une forme un peu différente qui la fera peut-être mieux comprendre.

([1]) Il serait très aisé de déduire de chacune de ces deux solutions des solutions de même nature et ayant respectivement les mêmes propriétés; mais ces généralisations aisées, qui peuvent peut-être rendre des services dans la pratique, ne paraissent pas avoir d'intérêt théorique.

On connaît, d'une part, des expressions analytiques qui ne peuvent représenter zéro sans être identiquement nulles et qui sont susceptibles de représenter des fonctions uniformes; mais, ou bien ces expressions analytiques ont un domaine de convergence qui diffère du domaine d'existence de la fonction (série de Taylor), ou bien elles ne peuvent représenter que des fonctions soumises à certaines conditions restrictives (théorème de M. Mittag-Leffler). On connaît, d'autre part, des expressions analytiques susceptibles de représenter des fonctions uniformes quelconques dans tout leur domaine d'existence; mais ces expressions peuvent représenter zéro sans être identiquement nulles (théorème de M. Runge).

Le but idéal à atteindre, c'est de trouver une représentation réunissant les avantages de la série de Taylor et des séries de M. Runge ou de M. Painlevé, sans avoir aucun de leurs inconvénients, et le but immédiat, c'est de trouver une telle représentation pour des classes de fonctions de plus en plus étendues.

Nous avons jusqu'ici utilisé la notion de fonction analytique sous la forme précise que lui a donnée Weierstrass; il n'est pas inutile de remarquer que la solution complète de notre problème conduirait probablement à une grande extension de cette notion. Ceci pourra être expliqué plus clairement après l'étude un peu détaillée de certaines expressions analytiques particulières.

Remarque sur les séries de fractions rationnelles.

Les expressions analytiques uniformes qui se présentent les premières sont les séries des fractions rationnelles (1). Soit

$$\sum \frac{P_n(z)}{Q_n(z)}$$

une telle série; $P_n(z)$ et $Q_n(z)$ sont des polynomes. Nous supposons, bien entendu, qu'il existe au moins une région du plan dans

(1) Elles comprennent comme cas particulier les séries de polynomes (et les séries de puissances); d'ailleurs, en changeant z en $\frac{1}{z-a}$ une série de polynomes devient une série de fractions rationnelles; elle ne s'en distinguait que par le rôle spécial du point à l'infini.

laquelle cette série est convergente; nous supposerons de plus qu'il existe une région dans laquelle la convergence est absolue et uniforme. Cela étant, deux cas sont à distinguer : la série peut rester, ou ne pas rester, absolument et uniformément convergente lorsqu'on décompose les fractions rationnelles en éléments simples. Nous exclurons le dernier cas : indiquons brièvement pour quelles raisons. Soit

$$\sum \frac{A_n}{(z-a_n)^{m_n}}$$

la série obtenue après la décomposition en éléments simples [1]; si cette série ne converge pas absolument, elle a des valeurs différentes lorsqu'on modifie l'ordre de ses termes; par suite, la fonction n'est pas déterminée par la connaissance de ces singularités (les points a_n) et de la manière dont elle se comporte en ces points (définie par les coefficients A_n et les exposants m_n). C'est là un premier motif pour ne pas considérer de telles séries. En voici un deuxième; nous serons amenés tout à l'heure à supposer que les valeurs des exposants m_n sont inférieures à un nombre fixe; sinon la série pourrait représenter zéro; il est aisé de voir que, même avec cette restriction, en supposant, par exemple, les m_n tous égaux à *un*, on aurait le même inconvénient si l'on ne suppose pas la série absolument convergente.

En effet, si l'on a une série de fractions rationnelles ayant des pôles multiples, on peut trouver une série de fractions rationnelles qui en diffère aussi peu qu'on veut (en excluant des régions aussi petites qu'on veut autour des pôles) et qui n'ait que des pôles simples. Cette remarque, que je dois à M. Painlevé, suffit pour démontrer que l'hypothèse de la convergence absolue est nécessaire si l'on veut que la série à pôles simples ne puisse pas représenter zéro. Mais nous n'insistons pas sur ces considérations, en somme accessoires, et seulement destinées à montrer que les restrictions que nous faisons ne sont pas purement arbitraires, mais découlent de la manière dont nous avons posé le problème de la représentation analytique des fonctions uniformes.

---

(1) On pourrait avoir, pour une ou plusieurs valeurs de n, $a_n = \infty$; $\frac{1}{z-a_n}$ devrait alors être remplacé par z.

Nous supposons donc que la série

$$\sum \frac{A_n}{(z-a_n)^{m_n}}$$

est absolument et uniformément convergente dans au moins une région du plan. Les points a_n, que nous appelons *les pôles de la série* ne sont pas nécessairement tous distincts; l'ordre d'un pôle a_n est le plus grand des exposants m_n qui lui correspondent. Nous exclurons le cas où cet ordre est infini; on peut, en effet, faire alors deux hypothèses très différentes. Il peut arriver, d'une part, que l'ensemble des termes qui correspondent à un même pôle a constitue une fonction entière de $\frac{1}{z-a}$; c'est ce qui a lieu dans les développements de M. Mittag-Leffler dont nous avons déjà parlé et que nous n'avons pas l'intention d'étudier ici.

Il peut arriver, d'autre part, que l'ensemble des termes qui correspondent au pôle a forme une série qui converge seulement à l'extérieur d'un certain cercle a. M. Appell a donné, dans un Mémoire déjà cité, des exemples de séries de ce genre, avec trois pôles, et représentant *zéro* dans une région du plan, et *un* dans une autre région. C'est une raison suffisante pour exclure ces séries de nos considérations.

Nous supposons donc non seulement que la série est absolument et uniformément convergente, mais encore que l'ordre de chaque pôle est fini. Là se bornent les hypothèses restrictives rendues indispensables par la nature du problème à résoudre [1]. Nous allons malheureusement être obligés de faire d'autres hypothèses, d'un caractère plus arbitraire, et justifiées seulement par le désir d'obtenir des résultats simples. Mais nous avons tenu à les distinguer nettement des précédentes [2].

---

[1] On pourrait sans doute montrer aisément qu'il est nécessaire, au même point de vue, de supposer que les ordres des divers pôles sont tous inférieurs à un nombre fixe : ce n'est pas une conséquence du fait que chacun de ces ordres est fini. Mais ce point nous semble très peu important.

[2] Les restrictions que nous allons introduire ne sont évidemment pas les seules qui conduisent à des résultats simples ; il est évidemment aisé d'en trouver d'autres; il semble qu'il le soit beaucoup moins de les trouver *toutes* et surtout de les étudier méthodiquement.

Étude de certaines séries de fractions simples.

Les hypothèses restrictives peuvent porter, soit sur les pôles a_n soit sur les coefficients A_n, soit enfin sur les uns et sur les autres. Nous allons ici, pour fixer les idées, ne faire *aucune hypothèse* sur les points a_n; leur distribution dans le plan pourra être absolument quelconque ([1]); mais nous serons amenés à imposer des restrictions considérables aux exposants m_n et aux coefficients A_n. On pourrait, au contraire, en faisant certaines restrictions sur la distribution des a_n, obtenir des résultats ([2]) avec moins de restrictions sur les A_n et les m_n. Mais notre but ici est moins d'atteindre une grande généralité que de montrer comment on peut obtenir des résultats précis avec des hypothèses, en somme assez larges. *Nous supposerons tous les exposants m_n égaux à l'unité;* nous considérons donc la série

$$\sum \frac{A_n}{z - a_n}$$

et nous supposons, en outre, qu'on a ([3])

$$\lim_{n=\infty} \sqrt[n]{A_n} = 0.$$

([1]) On suppose cependant qu'il existe des aires ne renfermant aucun point a_n; sinon il ne pourrait pas exister des aires *en tous les points* desquelles la série soit convergente.

([2]) *Voir*, par exemple, E. Borel, *Sur quelques points de la théorie des fonctions*, première Partie (*Annales de l'École Normale*, 1895).

([3]) Cette hypothèse est identique à la suivante; la série

$$\Sigma A_n Z^n$$

a un rayon de convergence infini; c'est une fonction entière de Z. On peut remarquer que nous avons été amenés par la considération de résultats de M. Appell, à faire une hypothèse analogue à propos des séries de la forme

$$\sum \left(\frac{A_n}{(z-a)^n} + \frac{B_n}{(z-b)^n} + \frac{C_n}{(z-c)^n} \right); \tag{1}$$

si aucune des trois séries

$$\Sigma A_n Z^n, \quad \Sigma B_n Z^n, \quad \Sigma C_n Z^n$$

n'avait un rayon de convergence infini, la série (1) pourrait représenter *zero*

Enfin, rappelons que l'ensemble des a_n est absolument quelconque, mais que cependant *il n'est pas dense dans tout le plan;* il existe, par suite, des aires ne renfermant pas de points a_n et dans chacune desquelles la série définit une fonction analytique. Mais, dans des régions de convergence séparées, les deux fonctions analytiques ainsi définies n'ont *a priori* aucune relation entre elles [1]; nous allons démontrer que si la somme de la série est identiquement nulle dans une région du plan, elle est nulle dans toute région où elle converge : ce qui est une des propriétés caractéristiques des séries de Taylor.

Pour démontrer cette proposition nous remarquerons d'abord que si l'on pose

$$z = \frac{\alpha z' + \beta}{\gamma z' + \delta} \qquad (\alpha\delta - \beta\gamma \neq 0),$$

$$a_n = \frac{\alpha a'_n + \beta}{\gamma a'_n + \delta},$$

on a

$$z - a_n = \frac{(\alpha\delta - \beta\gamma)(z' - a'_n)}{(\gamma a'_n + \delta)(\gamma z' + \delta)},$$

et

$$\frac{A_n}{z - a_n} = \frac{A'_n}{z' - a'_n} + B_n,$$

---

dans une de ses régions de convergence et *un* dans l'autre. Il est curieux qu'on soit amené à faire la même restriction sur les coefficients de la série

$$\sum \frac{A_n}{z - a_n}$$

que sur ceux de la série

$$\sum \frac{A_n}{(z - a)^n}.$$

[1] Si l'on a une région dans laquelle il n'y a pas de point a et que cette région soit limitée par une courbe fermée dont tous les points appartiennent à l'ensemble dérivé de l'ensemble des points a, on ne doit pas regarder comme évident que cette ligne est une *limite naturelle* de la fonction analytique égale à la série dans la région considérée. Ce point n'est même pas aisé à démontrer dans le cas général [on peut consulter à ce sujet le Mémoire déjà cité en Note (p. 93)]; mais il est facile de voir que les cas dans lesquels il n'en serait pas ainsi, doivent être regardés comme exceptionnels et il me paraît même possible de démontrer que ces cas ne peuvent se présenter avec les restrictions faites ici sur les A_n. Mais l'étude détaillée de ces questions exigerait de longs développements dont ce n'est pas ici le lieu. Nous tenons surtout à faire voir qu'on peut obtenir certains résultats *sans aucune hypothèse sur les a_n*.

en posant

$$A'_n = \frac{A_n(\gamma a'_n + \delta)^2}{\alpha\delta - \beta\gamma},$$

$$B_n = A_n \frac{\gamma(\gamma a'_n + \delta)}{\alpha\delta - \beta\gamma}.$$

On peut écrire enfin

$$(1) \qquad \sum \frac{A_n}{z - a_n} = B + \sum \frac{A'_n}{z' - a'_n},$$

en posant

$$B = \Sigma B_n.$$

L'égalité (1) exprime que *la fonction considérée conserve sa forme par une transformation homographique* [1]. Par suite, s'il existe deux régions séparées dans lesquelles la série soit convergente, si l'on désigne par c et d deux points appartenant respectivement à ces régions, il suffira de poser

$$z' = \frac{z - c}{z - d},$$

pour que la série transformée en z' soit convergente dans le voisinage des points $z' = 0$ et $z' = \infty$, c'est-à-dire à l'intérieur d'un cercle de rayon assez petit et à l'extérieur d'un cercle de rayon assez grand. D'ailleurs, entre ces deux cercles, la distribution des points a'_n est quelconque. Reprenant les anciennes notations, nous considérons une série de la forme

$$B + \sum \frac{A_n}{z - a_n},$$

les points a étant tous compris à l'intérieur d'une couronne circulaire ayant pour centre le point $z = 0$. Quant aux A, nous supposons simplement pour l'instant que la série $\Sigma |A_p|$ est convergente; nous introduirons seulement tout à l'heure les autres hypothèses que nous avons faites à leur sujet.

Nous nous proposons, étant admis que le développement en série de Taylor de la fonction proposée à l'intérieur du petit

---

(1) On voit aisément que les conditions de convergence supposées pour les A_n sont sûrement vérifiées pour les A'_n, si le point $z = \frac{\alpha}{\gamma}$ n'est pas un point a_n; c'est bien ce qui a lieu dans l'application que nous ferons dans un instant.

cercle de la couronne est identiquement nul [1], de chercher ce qu'on en peut conclure pour le développement à l'extérieur du grand cercle. Posons

$$\sum_{n+1} |A_p| = \varepsilon_n;$$

ε_n désigne l'erreur qu'on commet en s'arrêtant au $n^{\text{ième}}$ terme dans le calcul de la somme des modules des A.

Cela étant, nous avons, en désignant par x_n l'inverse des a_n,

$$\frac{1}{a_n - z} = \frac{x_n}{1 - x_n z} = \sum_{p=1}^{p=\infty} x_n^p z^{p-1}$$

et, par suite,

$$\sum \frac{A_n}{a_n - z} = \sum_{n=1}^{n=\infty} \sum_{p=1}^{p=\infty} A_n x_n^p z^{p-1}.$$

Nous avons donc, par hypothèse [2], quel que soit p,

$$(1) \qquad \Sigma A_n x_n^p = 0 \qquad (p = 1, 2, 3, \ldots).$$

Désignons par $\frac{1}{\alpha}$ et $\frac{1}{\beta}$ les rayons des deux cercles qui limitent la couronne; on a

$$\beta < |x_n| < \alpha$$

et, par suite,

$$\left| \sum_{h=n+1}^{h=\infty} A_h x_h^p \right| < \varepsilon_n \alpha^p.$$

Nous avons donc, θ_p ayant un module inférieur à l'unité,

$$\begin{aligned}
A_1 x_1 + A_2 x_2 + \ldots + A_n x_n &= \theta_1 \alpha \varepsilon_n, \\
A_1 x_1^2 + A_2 x_2^2 + \ldots + A_n x_n^2 &= \theta_2 \alpha^2 \varepsilon_n, \\
\ldots\ldots\ldots\ldots&\ldots\ldots\ldots\ldots, \\
A_1 x_1^n + A_2 x_2^n + \ldots + A_n x_n^n &= \theta_n \alpha^n \varepsilon_n.
\end{aligned}$$

Posons

$$(x - x_1)(x - x_2)\ldots(x - x_h) = x^n + u_1 x^{n-1} + u_2 x^{n-2} + \ldots + u_n.$$

---

[1] C'est évidemment une conséquence immédiate du fait que la somme de la série est nulle dans une région intérieure au petit cercle.

[2] Dans le cas où B n'est pas nul, la relation n'est pas vérifiée pour $p = 1$, mais on voit aisément que ce point n'a aucune importance.

Nous conclurons facilement de ce qui précède qu'en désignant par σ_n la somme $A_1 + A_2 + \ldots + A_n$ on a

$$\varepsilon_n[\theta_n \alpha^n + u_1 \theta_{n-1} \alpha^{n-1} + u_2 \theta_{n-2} \alpha^{n-2} + u_3 \theta_{n-3} \alpha^{n-3} + \ldots + u_{n-1} \theta_1 \alpha] + u_n \sigma_n = 0.$$

Or on a visiblement

$$|u_n| > \beta^n,$$
$$|\theta_n \alpha^n + u_1 \theta_{n-1} \alpha^{n-1} + \ldots + u_{n-1} \theta_1 \alpha| < \alpha^n + |u_1| \alpha^{n-1} + \ldots + |u_{n-1}| \alpha + |u_n|.$$

Or le second membre de cette inégalité est évidemment inférieur à

$$(\alpha + |x_1|)(\alpha + |x_2|) \ldots (\alpha + |x_n|) < (2\alpha)^n.$$

Nous avons donc finalement

$$|\sigma_n| < \varepsilon_n \left(\frac{2\alpha}{\beta}\right)^n.$$

Donc, si nous supposons qu'on ait

$$(2) \qquad \lim_{n=\infty} \varepsilon_n \left(\frac{2\alpha}{\beta}\right)^n = 0,$$

la série $A_1 + A_2 + \ldots + A_n + \ldots$ aura pour somme zéro. Dans ces conditions nous démontrerons absolument de même que la série

$$\frac{A_1}{x_1} + \frac{A_2}{x_2} + \ldots + \frac{A_n}{x_n} + \ldots$$

a pour somme zéro; il suffit, en effet, de poser $\dfrac{A_p}{x_p} = A'_p$ et de raisonner sur les A' comme nous venons de le faire sur les A, les ε seront multipliés par un facteur compris entre α et β, et la condition (2) continuera à être vérifiée. Nous démontrerons donc que, quel que soit h, on a

$$\frac{A_1}{x_1^h} + \frac{A_2}{x_2^h} + \ldots + \frac{A_n}{x_n^h} + \ldots = 0.$$

Il en résulte que le développement de la série

$$\sum \frac{A_n}{a_n - z},$$

suivant les puissances décroissantes de z, développement valable pour les points extérieurs à la couronne considérée, est identiquement nul.

Il est clair que la condition (1) est vérifiée si le rayon de convergence de la série

$$\Sigma A_n Z_n$$

est supérieur à $\frac{2\alpha}{\beta}$. Or nous avons supposé cette série convergente dans tout le plan de la variable Z. Il est alors aisé de voir que la fonction

$$\varphi(z) = B + \sum \frac{A_n}{z - a_n}$$

ne peut pas être nulle dans une portion du plan où elle est holomorphe sans être nulle dans tout le plan. En effet, soient C_1 et C_2 deux contours à l'intérieur desquels elle est holomorphe, et C' un contour assez grand pour ne renfermer à son extérieur aucun point a. Nous tracerons deux cercles concentriques, l'un extérieur à C_1, l'autre extérieur à C' et deux autres cercles concentriques, l'un intérieur à C_2, l'autre extérieur à C'. D'après ce que nous avons vu, la fonction supposée nulle à l'intérieur de C_1, par exemple, le sera à l'extérieur de C' et, par suite, à l'intérieur de C_2; ce qu'il fallait établir.

Propriétés essentielles des séries étudiées. — Conclusion.

Ces resultats s'étendent sans difficulté au cas où les termes de la série sont d'un même degré différent du premier; il suffit, en effet, d'une intégration ou de plusieurs intégrations successives pour rentrer dans le cas que nous venons de traiter; on voit aisément que les constantes introduites par l'intégration sont nulles, car on pourrait évidemment conclure de ce qui précède que la série $\varphi(z)$ ne peut pas être égale à un polynome en z. Cela revient, en effet, à supposer que les relations (1) ne sont vérifiées qu'à partir d'une certaine valeur de p; la démonstration en est à peine modifiée.

Nous avons tenu à donner cette démonstration parce qu'elle donne un exemple des procédés qu'on peut employer pour étu-

dier la série considérée lorsque les a sont quelconques; mais la proposition elle-même que nous avons démontrée nous paraît beaucoup moins importante que la suivante :

On peut tracer dans toutes les régions du plan une infinité non dénombrable de courbes sur lesquelles la série est absolument et uniformément convergente, ainsi que toutes ses dérivées.

Cette proposition est d'ailleurs une conséquence immédiate des théorèmes du Chapitre précédent [1]; nous renvoyons à ce Chapitre pour ce qui concerne l'*infinité* des courbes considérées.

Remarquons simplement que, sur chacune de ces courbes C, la série représente une fonction continue, admettant des dérivées de tous les ordres, qui sont représentées par les séries dérivées. Il en

---

[1] En effet, les modules des termes des diverses dérivées de la série considérée sont de la forme

$$\frac{|A_n|}{[(x-\alpha_n)^2+(y-\beta_n)^2]^{\frac{p}{2}}}$$

pour la dérivée d'ordre $p-1$, en posant $a_n=\alpha_n+i\beta_n$. D'ailleurs, l'hypothèse

$$\lim_{n=\infty} \sqrt[n]{A_n}=0$$

peut s'écrire

$$\lim_{n=\infty} \sqrt[n]{\frac{1}{|A_n|}}=\infty,$$

et, par suite, le logarithme ayant un sens arithmétique

$$\lim_{n=\infty} \frac{1}{n}\log\frac{1}{|A_n|}=\infty,$$

c'est-à-dire qu'on a, à partir d'une certaine valeur de n,

$$\log\frac{1}{|A_n|}>n,$$

et la série

$$\sum \frac{1}{\left[\log\frac{1}{|A_n|}\right]^2}$$

est convergente, car elle a ses termes, au moins à partir d'un certain rang, respectivement plus petits que ceux de la série convergente

$$\sum \frac{1}{n^2}.$$

résulte qu'en un point où se croisent des courbes C, en nombre fini ou infini, la dérivée de la fonction est la même, quelle que soit la courbe C sur laquelle on se déplace.

Citons une autre conséquence du fait que la série

$$\sum \frac{A_n}{z - a_n}$$

est absolument et uniformément convergente sur une courbe C.

On peut, le long de la courbe C, intégrer la série terme à terme et son intégrale est absolument et uniformément convergente et représente l'intégrale de la fonction. On en conclut le théorème suivant, qu'on peut considérer comme presque évident, mais qui ne semble pas aisé à démontrer directement : « Si l'ensemble des a est dense à l'intérieur d'une aire, il est impossible que, dans un espace E aussi petit qu'on veut de cette aire, les valeurs que prend la fonction sur les diverses courbes C qu'on peut tracer dans cet espace E coïncident avec les valeurs d'une fonction holomorphe sur ces mêmes courbes C ». En effet, l'intégrale de la fonction le long d'un contour fermé quelconque compris à l'intérieur de cette aire serait nulle; on en conclurait que la somme des résidus A_n de la fonction relatifs aux pôles compris à l'intérieur de ce contour serait nulle. Or, on peut tracer un contour tel que C, aussi voisin qu'on veut d'un contour quelconque donné à l'avance; de ce fait et de celui que la série des modules des A est convergente, on conclut immédiatement que tous les résidus sont nuls, c'est-à-dire que la série est identiquement nulle.

Mais nous n'insistons pas sur ces propriétés que l'on pourrait multiplier; nous voulions seulement montrer comment on pouvait introduire dans le calcul des expressions analytiques dont les valeurs, en des régions diverses de convergence, soient liées simplement entre elles. Il semble, dès lors, qu'on puisse songer à étendre la définition donnée par Weierstrass de la *fonction analytique* et regarder dans certains cas, comme étant *la même fonction*, des fonctions analytiques ayant des domaines d'existence séparés. Mais il nécessaire pour cela de faire des restrictions sur la nature des expressions analytiques que l'on considère, et c'est parce qu'il n'a voulu faire aucune restriction de ce genre que Weierstrass a résolu par la négative la question qu'il avait

posée sous la forme suivante : *Es war daher der Gedanke nicht abzuweisen, ob nicht überhaupt in dem Falle, wo ein arithmetischer Ausdruck* F(x) *in verschiedenen Teilen seines Geltungsbereichs verschiedene monogene Functionen der complexen Veränderlichen x darstellt, unter diesen ein nothwendiger Zusammenhang bestehe, der bewirke, dass durch die Eigenschaften der einen auch die Eigenschaften der andern bestimmt seien. Wäre dies der Fall, so würde es daraus folgen, dass der Begriff der monogenen Function erweitert werden müsste.* (WEIERSTRASS, *Werke*, t. II, p. 212.)

Il serait nécessaire de parler aussi d'une objection faite par M. Poincaré aux tentatives de prolongement d'une fonction au delà d'une ligne singulière fermée; nous renverrons pour ce point au Mémoire déjà cité en note (p. 93). Contentons-nous de dire que les arguments de M. Poincaré sont irréprochables, si l'on ne fait aucune restriction sur la nature des expressions analytiques qu'on considère; or nous avons insisté à diverses reprises sur la nécessité qu'il y avait, à notre point de vue, à faire de telles restrictions. Notre désaccord avec M. Poincaré n'est donc qu'apparent; il provient de ce que nous avons étudié deux problèmes, en apparence identiques, mais en réalité différents.

Il ne nous est pas possible de donner à ce Chapitre une conclusion décisive et définitive; car, à notre avis, la question qui y est examinée n'est pas entièrement résolue et appelle de nouvelles recherches. Nous serions satisfait si nous avions convaincu nos lecteurs que, ni les travaux fondamentaux de Weierstrass, ni les travaux ultérieurs de MM. Mittag-Leffler, Appell, Poincaré, Runge, Painlevé, ne résolvent entièrement la question des rapports entre la notion de fonction analytique et d'expression analytique. On peut même dire, sans exagération, que l'étude et la classification des expressions analytiques, non susceptibles de représenter zéro sans être identiquement nulles, est encore à faire à peu près entièrement [1].

---

[1] Pour le développement des idées de ce Chapitre, *voir* mes *Leçons sur les fonctions monogènes uniformes d'une variable complexe* (Gauthier-Villars, 1915). (*Note de la deuxième édition.*)

NOTE I.

LA NOTION DES PUISSANCES.

I. — L'égalité et l'inégalité des puissance.

Nous avons dit que deux ensembles ont *même puissance*, lorsqu'on peut établir entre leurs éléments une correspondance univoque et réciproque ; mais nous nous sommes contentés de considérer, d'une part, des ensembles dénombrables ou de *première puissance ;* d'autre part, des ensembles ayant la *puissance du continu;* c'est-à-dire que nous avons toujours comparé les puissances à des puissances *données* (ou considérées comme telles). Dans l'étude sommaire que nous allons faire de la notion générale de puissance, nous procéderons de même, en ce sens que nous ne séparerons jamais la notion abstraite de puissance de la notion plus concrète d'ensemble et que nous ne chercherons pas à concevoir *la puissance en soi*, indépendamment de tout ensemble possédant cette puissance [1].

Étant donnés deux ensembles A et B, nous désignerons par A_1 une partie aliquote quelconque de A, c'est-à-dire un ensemble comprenant uniquement des éléments de A, mais ne les comprenant pas tous; de même B_1 désignera une partie aliquote quelconque de B. Cela étant, si l'on compare A et B, quatre cas sont logiquement possibles et s'excluent réciproquement :

1° *Il existe un* A_1 *ayant même puissance que* B, *et il n'existe pas de* B_1 *ayant même puissance que* A.

2° *Il n'existe pas de* A_1 *ayant même puissance que* B, *et il existe un* B_1 *ayant même puissance que* A.

3° *Il existe un* A_1 *ayant même puissance que* B *et aussi un* B_1 *ayant même puissance que* A.

4° *Il n'existe ni un* A_1 *ayant même puissance que* B, *ni un* B_1 *ayant même puissance que* A.

---

[1] D'après M. G. Cantor, la notion de puissance s'obtient en partant de l'idée d'ensemble, par une *double* abstraction : abstraction de *l'ordre* et de la *nature* des objets. Mais il ne nous semble pas que cette remarque suffise pour permettre de raisonner sur les puissances indépendamment de tout substratum.

Les deux premiers cas ne diffèrent que par l'échange de A et de B; dans le premier cas, on dit que la puissance de A est supérieure à celle de B et dans le second que la puissance de B est supérieure à celle de A. On démontre d'ailleurs très aisément(1) que si la puissance d'un ensemble C est égale ou inférieure à celle d'un ensemble B, laquelle est égale ou inférieure à celle d'un ensemble A, la puissance de C est égale ou inférieure à celle de A, l'égalité n'ayant lieu que si les puissances de A et de C sont toutes deux égales à celles de B.

D'autre part, il est clair que ces deux premiers cas, non seulement sont logiquement possibles, mais sont réellement possibles; il suffit de prendre, pour l'un des ensembles, un ensemble dénombrable et, pour l'autre, un ensemble ayant la puissance du continu.

On peut aussi donner des exemples du troisième et du quatrième cas. Pour le troisième, il suffit de prendre, pour A et B, *deux ensembles infinis ayant même puissance;* nous savons, en effet (p. 13), qu'on peut, en retranchant de A, par exemple, un ensemble dénombrable, obtenir un ensemble A_1, qui est une partie aliquote de A et qui a même puissance que A, c'est-à-dire même puissance que B. On obtiendrait de même un ensemble B_1 ayant même puissance que B.

Le quatrième cas, enfin, se présente lorsqu'on a deux ensembles composés *d'un même nombre fini d'éléments*. Il est clair, en effet, qu'une partie aliquote quelconque de l'un deux ne saurait avoir même puissance que l'autre.

La question qui se pose maintenant est la suivante :

Lorsqu'on est dans l'un des deux derniers cas, peut-on affirmer que les deux ensembles A *et* B *ont même puissance* (2)?

Nous allons démontrer qu'il en est ainsi dans le troisième cas; mais, dans le quatrième cas, nous ne savons rien. C'est une question qu'il serait très important de résoudre; car, si on la résout par l'affirmative, dans le quatrième cas comme dans le troisième, on pourra affirmer que, deux ensembles quelconques étant donnés, ou bien ils ont même puissance, ou bien la puissance de l'un est supérieure à celle de l'autre, au sens precis que nous avons donné à ces mots. Si, au contraire, ce quatrième cas se résout par la négative, il pourra exister deux ensembles A et B tels qu'il soit impossible de *comparer leurs puissances :* la puissance de A ne pourra être dite ni égale, ni supérieure, ni inférieure à celle de B (3).

---

(1) Il suffit de se servir du procédé de *projection* que nous expliquons à la page suivante.

(2) D'ailleurs si l'on admet ce point, on prouvera aisément que, dans le troisième cas, les ensembles renferment une infinité d'éléments et dans le quatrième un nombre fini.

(3) Ces difficultés, dont j'avais vainement cherché la solution dans les publications de M. G. Cantor, m'avaient, à diverses reprises, vivement préoccupé, et

Tout en souhaitant vivement que cette question importante s'éclaircisse, il nous paraît bien difficile qu'on puisse tirer une démonstration précise et rigoureuse des hypothèses *purement négatives* du quatrième cas, si on laisse à l'idée d'ensemble toute sa généralité. C'est pour cette raison que nous n'avons pas cru devoir donner de place ici aux travaux de M. C. Cantor sur la *numération* des puissances, car il nous a semblé qu'on ne pouvait actuellement bâtir une telle numération sur une base absolument sûre. Mais nous sommes loin de méconnaître l'importance de ces recherches, car, en dehors de leur intérêt philosophique et même si elles n'aboutissent pas à leur but idéal, elles peuvent rendre de très grands services. Par exemple, nous aurons à nous servir bientôt (p. 127) du théorème que nous allons maintenant démontrer : *Dans le troisième cas les puissances de* A *et de* B *sont égales.* Rappelons que les hypothèses, dans ce troisième cas, sont les suivantes :

Il existe une partie aliquote A_1, de A, qui a même puissance que B et il existe une partie aliquote B_1, de B, qui a même puissance que A. Il faut prouver que A a même puissance que B.

Les ensembles B et A_1 ayant même puissance, il existe une *projection* de B sur A_1, c'est-à-dire une loi d'après laquelle les éléments de B et de A_1 se correspondent d'une manière univoque et réciproque. Il existe même une infinité de telles projections; mais nous en choisissons une bien déterminée. Il est clair que, par une telle projection, à toute partie aliquote de B correspond une partie aliquote de A_1; soit A_2 la partie aliquote de A_1 qui correspond ainsi à B_1; A_2 a, d'après la définition même de la puissance, même puissance que B_1 et, par suite, même puissance que A. D'ailleurs, A_2 *est une partie aliquote de* A_1, *qui est lui-même une partie aliquote de* A. Tout revient à dire que, A_2 *ayant même puissance que* A, A_1 *a aussi même puissance que* A (car B a même puissance que A_1).

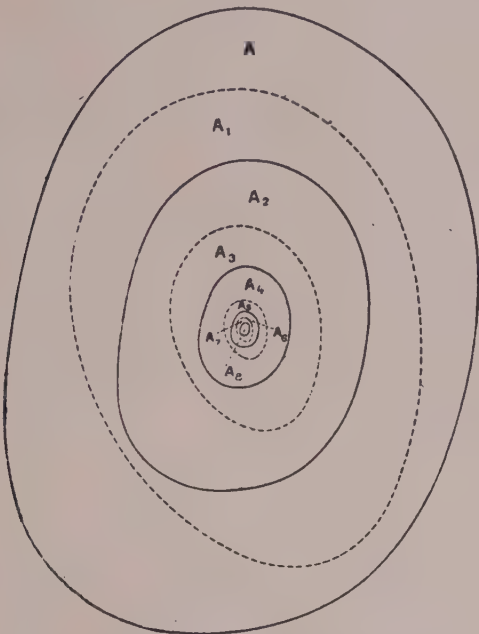
Par hypothèse A_2 a même puissance que A; choisissons une projection déterminée de A sur A_2; A_1 qui est une partie aliquote de A deviendra une partie aliquote A_3 de A_2, et A_2 qui est une partie aliquote de A_1 deviendra une partie aliquote A_4 de A_3. C'est ce que nous indiquons sur la figure

---

ayant eu l'honneur de faire la connaissance de M. G. Cantor au Congrès de Zurich (août 1897), je m'empressai de les lui soumettre. Il voulut bien m'indiquer, pour le troisième cas, le théorème que nous venons d'énoncer et dont la démonstration termine ce paragraphe; cette démonstration inédite est due à M. Félix Bernstein et a été donnée pour la première fois dans le séminaire de M. G. Cantor, à Halle. Quant au théorème lui-même, M. G. Cantor le considère comme exact depuis fort longtemps; il a bien voulu me communiquer verbalement que son sentiment est le même au sujet du quatrième cas; j'indique dans le texte les raisons qui ne me permettent pas de partager entièrement sa manière de voir sur ce dernier point. Qu'il me soit permis de remercier vivement M. G. Cantor de l'obligeance avec laquelle il m'a communiqué les renseignements qui précèdent et m'a autorisé à les publier ici.

schématique ci-dessous, où la *projection* de A sur A_2 est une transformation homothétique. D'ailleurs A_3 a même puissance que A_1 et A_4 même puissance que A_2 et que A. Si nous *projetons* A sur A_4, A_1 et A_2 seront projetés suivant A_5 et A_6; A_5 sera une partie aliquote de A_4 et aura même

Fig. 2.



puissance que A_1 et A_3; A_6 sera une partie aliquote de A_5 et aura même puissance que A_1, A_2, A_4. En continuant ainsi, on formera une suite d'ensembles

$$A, \quad A_1, \quad A_2, \quad A_3, \quad A_4, \quad \ldots,$$

tels que chacun d'eux est une partie aliquote du précédent et tels de plus que tous les ensembles d'indice pair ont même puissance que A et tous les ensembles d'indice impair même puissance que A_1.

D'ailleurs cette suite se prolonge indéfiniment, c'est-à-dire que l'ensemble A_n est défini, quel que soit l'entier n; il renferme d'ailleurs sûrement des éléments, puisqu'il a même puissance que A ou que A_1, suivant que n est pair ou impair.

Considérons maintenant ([1]) l'ensemble D formé des éléments communs

([1]) En introduisant l'ensemble D, je modifie légèrement la démonstration que M. G. Cantor m'avait communiquée (*voir* Note de la page 104), afin d'éviter l'introduction des nombres transfinis. Cette modification est d'ailleurs sans importance.

à tous les ensembles $A, A_1, A_2, \ldots, A_n, \ldots$. Cet ensemble D peut d'ailleurs ne renfermer aucun élément. Il est clair que cet ensemble D peut s'obtenir en retranchant successivement de A les ensembles

$$A - A_1, \quad A_1 - A_2, \quad A_2 - A_3, \quad \ldots, \quad A_n - A_{n+1}, \quad \ldots.$$

Nous pouvons donc écrire

$$A = D + (A - A_1) + (A_1 - A_2) + (A_2 - A_3) + \ldots$$

et chacun des symboles entre parenthèses désigne un ensemble déterminé, puisque A_{n+1} est une partie aliquote de A_n. On peut écrire de même

$$A_1 = D + (A_1 - A_2) + (A_2 - A_3) + (A_3 - A_4) + \ldots.$$

Il est maintenant aisé de démontrer que les ensembles A et A_1 ont même puissance; il suffit de remarquer qu'on peut les regarder comme formés d'une infinité dénombrable d'ensembles ayant deux à deux même puissance. Il résulte, en effet, du mode de projection par lequel on a obtenu les ensembles $A_3, A_4, \ldots$ que $A - A_1$ a même puissance que $A_2 - A_3$, que $A_4 - A_5$, que $A_6 - A_7, \ldots$. De même, les ensembles $A_1 - A_2$, $A_3 - A_4$, $A_5 - A_6, \ldots$ ont tous même puissance. Il suffit, dès lors, d'écrire l'expression de A_1 sous la forme

$$A_1 = D + (A_2 - A_3) + (A_1 - A_2) + (A_4 - A_5) + (A_3 - A_4) + (A_6 - A_7) + \ldots,$$

pour reconnaître que chacun des termes de cette série a même puissance que le terme de même rang de la série qui définit A.

Le théorème est donc démontré [1]. Nous en ferons surtout usage sous la forme suivante : Lorsqu'une partie aliquote de A aura même puissance que B, nous dirons que *la puissance de* A *est inférieure ou égale à celle de* B *et que la puissance de* B *est supérieure ou égale à celle de* A. Dès lors, lorsque nous aurons pu démontrer que la puissance d'un ensemble A est *à la fois supérieure ou égale* et *inférieure ou égale* à celle d'un

---

[1] Il importe de signaler une grave erreur qui pourrait être suggérée par la démonstration précédente. Nous avons vu que $A - A_1$ et $A_2 - A_3$ avaient même puissance, à cause de la projection au moyen de laquelle nous avions obtenu A_3; *mais ce n'est nullement une conséquence du fait que* A_2 *a même puissance que* A *et* A_3 *même puissance que* A_1. Il est aisé de s'assurer que, deux ensembles A et B ayant même puissance, et leurs parties aliquotes respectives A_1 et B_1 ayant aussi même puissance, *on n'en peut rien conclure pour* $A - A_1$ *et* $B - B_1$. Il suffit de prendre, par exemple, pour A l'ensemble des points compris entre 0 et 2, pour B l'ensemble des points compris entre 0 et 1, et pour A_1, ainsi que pour B_1, l'ensemble des points d'abscisses incommensurables compris entre 0 et 1. L'ensemble $A - A_1$ a alors la puissance du continu, tandis que l'ensemble $B - B_1$ est dénombrable. On aurait pu d'ailleurs supposer aussi A_1 et B_1 tous deux identiques à B.

ensemble B, nous pourrons affirmer qu'*elle lui est égale*. Sous cette forme, le théorème paraît évident; mais il importe de remarquer que s'il n'avait pas été préalablement démontré, il n'eût pas été légitime d'introduire ce langage.

II. — La formation d'ensembles ayant des puissances de plus en plus grandes.

On pourrait se demander si tous les ensembles n'ont pas, ou bien la puissance des ensembles dénombrables, ou bien la puissance du continu et, s'il en était ainsi, l'utilité pratique, sinon l'intérêt théorique des recherches sur les puissances en général serait singulièrement diminuée. Nous allons voir qu'on peut, par une méthode due à M. G. Cantor, *définir* des ensembles ayant des puissances de plus en plus grandes; nous rechercherons ensuite, à la fin de ce paragraphe, si l'on peut *concevoir* de tels ensembles.

Indiquons d'abord comment on peut définir un ensemble ayant une puissance supérieure à celle d'un ensemble donné E. Désignons par x un élément quelconque de E et par $f(x)$ une fonction de l'élément x, *ne pouvant prendre que l'une des deux valeurs* o *et* 1. Cette fonction est d'ailleurs définie d'une manière déterminée pour chaque élément de E. Je dis que l'ensemble F formé par *toutes* les fonctions telles que $f(x)$ a une puissance supérieure à la puissance de E. Bien entendu, deux fonctions $f(x)$ sont regardées comme distinctes si elles n'ont pas la même valeur pour *tous* les éléments de E.

Montrons d'abord qu'on peut trouver une partie aliquote F_1 de F ayant même puissance que E. Il suffit de faire correspondre à chaque élément x de E la fonction $f(x)$ qui prend la valeur o pour cet élément x et la valeur 1 pour tous les autres éléments de E. L'ensemble de ces fonctions $f(x)$, qui sont toutes distinctes, constitue une partie aliquote F_1 qui a même puissance que E. Donc la puissance de F est supérieure ou égale à celle de E; il nous suffit de faire voir qu'elle ne lui est pas égale, pour que nous soyons assurés qu'elle lui est supérieure. C'est là une conséquence des définitions et théorèmes du paragraphe précédent.

Si l'ensemble F a une même puissance que l'ensemble E, nous pouvons désigner par $f_y(x)$ la fonction qui correspond à un élément y de F et, si y est un élément quelconque de E, le symbole $f_y(x)$ pourra représenter un élément quelconque de F. Or, nous allons faire voir qu'il existe une fonction $f(x)$ distincte de toutes les fonctions $f_y(x)$; il suffit de considérer la fonction $f(x)$ définie par les inégalités

$$f(x) \neq f_x(x),$$

dans lesquelles x est un élément quelconque de E. Ces inégalités définissent complètement la fonction $f(x)$, car sa valeur pour un élément

quelconque x de E ne pouvant être que o ou 1, elle est connue, si l'on sait qu'elle n'est pas égale à $f_x(x)$ qui ne peut être aussi que o ou 1. Or, il est clair que la fonction $f(x)$ ainsi définie n'est égale à aucune des fonctions $f_y(x)$; car, si $f_y(x)$ est l'une de ces fonctions, la fonction $f(x)$ ne prend pas la même valeur pour $x = y$, car on a

$$f(y) \neq f_y(y)$$

Donc, la puissance de F n'est pas égale à la puissance de E; donc elle lui est supérieure.

Il est aisé de voir que, si l'on prend pour ensemble E un ensemble dénombrable, l'ensemble F a la puissance du continu; car, si à chacun des nombres 0, 1, 2, ..., n, ... on fait correspondre la valeur 0 ou 1, on peut représenter cette correspondance par une fraction « décimale »

$$0,10111001\ldots$$

supposée écrite dans le système de numération binaire. Nos lecteurs achèveront aisément le raisonnement à l'aide des principes posés dans le Chapitre I. Ils reconnaîtront, en particulier, que le cas où, à partir d'un certain rang, tous les chiffres « décimaux » seraient égaux à 1, donne lieu à une difficulté aisée à écarter.

Si l'ensemble E a la puissance du continu, on peut dire que l'ensemble F est, par exemple, l'ensemble des fonctions discontinues d'une variable réelle, prenant seulement l'une des deux valeurs 0 ou 1. Cet ensemble a une puissance supérieure à celle du continu [1].

Considérons cet ensemble F et séparons toutes les fonctions qu'il renferme en deux classes, cette séparation étant d'ailleurs faite suivant un mode quelconque.

L'ensemble de ces modes de séparation a une puissance supérieure à l'ensemble F. On peut continuer ainsi indéfiniment, c'est-à-dire construire une infinité dénombrable d'ensembles ayant des puissances de plus en plus grandes. D'ailleurs, lorsqu'on a une telle infinité d'ensembles

$$E_1, \quad E_2, \quad \ldots, \quad E_n, \quad \ldots,$$

tels que la puissance de chacun est supérieure à la puissance du précédent, il est clair que l'ensemble formé par la réunion de tous leurs éléments a une puissance supérieure à celle de chacun d'eux. Cette remarque, qui est due à M. G. Cantor, montre combien est considérable la multiplicité des puissances possibles; mais nous n'insistons pas sur ce point. Nous préférons consacrer la fin de ce paragraphe à la discussion du procédé de

---

[1] On verrait d'ailleurs aisément que l'on aurait un ensemble de même puissance que F, si l'on supprimait la restriction que $f(x)$ est nécessairement égal à 0 ou à 1, et que l'on considère l'ensemble des fonctions discontinues quelconques. (*Voir*, pour ce point, page 124.)

M. G. Cantor, dans le cas le plus simple où il nous fait connaître un ensemble dont la puissance est supérieure à celle du continu : l'ensemble F des fonctions discontinues, de variable réelle, ne prenant que les valeurs 0 et 1. Cet ensemble est logiquement défini; mais je me demande si nous en avons quelque conception. Pouvons-nous, en effet, concevoir la fonction discontinue la plus générale d'une variable réelle (même en supposant que les seules valeurs de la fonction sont 0 ou 1)? Il est nécessaire, en effet, pour donner une telle fonction de donner sa valeur pour *toutes* les valeurs réelles de la variable. Or, cet ensemble de valeurs n'étant pas dénombrable, il n'est pas possible d'indiquer un procédé qui permette de les avoir toutes, c'est-à-dire d'en atteindre une quelconque au bout d'un temps limité. On voit quelle est la difficulté qui se présente ici; elle est autrement grave que la difficulté analogue dans la définition de l'incommensurable générale, pour laquelle nous renverrons à un article de M. J. Tannery déjà cité (en note au bas de la page 3). En effet, si nous ne pouvons pas définir l'incommensurable générale de la même manière qu'un ensemble dénombrable, nous pouvons, du moins, semble-t-il, concevoir comme défini un nombre incommensurable quelconque, car nous pouvons définir les n premiers chiffres d'une fraction décimale, quel que soit n, et faire croître n indéfiniment. Nous ne pouvons rien faire de pareil pour la fonction discontinue générale d'une variable réelle, car il faut donner sa valeur *pour une infinité non dénombrable de valeurs de la variable*. Nous reparlerons plus loin de cette difficulté.

III. — La puissance des ensembles de fonctions.

Aussi nous semble-t-il difficile d'introduire la considération de telles fonctions et, en général, de toutes les classes de fonctions dont l'ensemble a une puissance supérieure à celle du continu; car, on ne peut se servir, dans le calcul, d'une fonction que si elle est définie au moyen d'une infinité dénombrable d'éléments. Tel est le cas des fonctions continues, et aussi des fonctions discontinues seulement pour une infinité dénombrable de valeurs de la variable. De telles fonctions peuvent, en effet, être définies, au moyen d'une infinité dénombrable de conditions et même, si l'on veut préciser davantage, au moyen d'une infinité dénombrable de nombres entiers [1].

Il en est de même pour les ensembles de points. On peut dire que chaque fonction $f(x)$, égale à 0 ou à 1 pour toute valeur réelle de la variable, sépare les nombres réels en deux classes et, par suite, définit deux ensembles de points : l'ensemble des points pour lesquels la fonction a la valeur 0 et l'ensemble des points pour lesquels la fonction a la valeur 1. Donc, l'ensemble F de ces fonctions est identique à l'ensemble de tous les

---

[1] Ces conditions ne sont d'ailleurs pas arbitraires; nous reviendrons sur ce point dans la Note III.

ensembles possibles (ayant pour éléments des nombres réels); cet ensemble a donc une puissance supérieure à celle du continu et son élément *général* (c'est-à-dire un ensemble *arbitraire* de nombres réels) ne peut être défini qu'au moyen d'un ensemble de conditions ayant la puissance du continu. C'est ce qui rend les raisonnements sur ces ensembles généraux si difficiles et parfois impossibles; c'est ce qui les rend aussi le plus souvent inutiles dans l'état actuel de la Science, car il ne semble pas qu'il soit aisé d'introduire dans un raisonnement l'un de ces ensembles, défini par une infinité non dénombrable de conditions. Au contraire, nous avons vu (p. 50) que l'ensemble des ensembles parfaits a la puissance du continu. On voit aisément qu'il en est de même des ensembles que nous avons appelés *mesurables;* c'est ce qui permet de raisonner plus aisément sur ces classes particulières d'ensembles et d'en découvrir des propriétés.

En résumé, il y a lieu, soit dans les ensembles de points, soit dans les fonctions discontinues, de distinguer deux grandes classes : les ensembles et les fonctions qui ne peuvent pas être définis par une infinité dénombrable de conditions et ceux qui peuvent l'être. Ces derniers seuls semblent pouvoir être actuellement considérés avec profit.

NOTE II.

LA CROISSANCE DES FONCTIONS ET LES NOMBRES DE LA DEUXIÈME CLASSE.

I. — Le théorème de Paul du Bois-Reymond.

On sait que, pour mesurer une longueur L avec une unité de longueur donnée l, on remarque qu'on peut trouver un entier n tel qu'on ait

$$nl \leqq \mathrm{L} < (n+1)l.$$

Si l'on pose $\mathrm{L} - nl = l_1$, on pourra trouver un entier n_1 non nul et tel que

$$n_1 l_1 \leqq l \leqq (n_1+1) l_1.$$

En continuant de même, on posera $l - n_1 l_1 = l_2$ et l'on définira la mesure de L par la fraction continue

$$n + \cfrac{1}{n_1 + \cfrac{1}{n_2 + \ldots}}$$

Ce procédé reçoit souvent le nom d'*algorithme d'Euclide;* on sait que son succès tient aux faits suivants :

1° On a la notion de la suite indéfinie des nombres entiers;

2° Etant données deux grandeurs de même nature L et l, il existe un nombre entier m tel que $ml > \mathrm{L}$: c'est ce qu'on appelle, en général, l'*axiome d'Archimède;*

3° Enfin, tout nombre incommensurable peut être entièrement défini par une infinité dénombrable de nombres commensurables.

Considérons deux fonctions $f(x)$ et $g(x)$ de la variable réelle positive x et supposons que ces fonctions soient croissantes et qu'elles croissent indéfiniment lorsque x augmente indéfiniment; supposons, de plus, que le rapport $\dfrac{f(x)}{g(x)}$ croisse indéfiniment avec x; nous pourrons convenir d'écrire

$$f(x) > g(x)$$

et de dire que la croissance de $f(x)$ est supérieure à la croissance de $g(x)$.

Si $\frac{f(x)}{g(x)}$ tend vers une limite finie (ou reste compris entre des limites finies) lorsque x augmente indéfiniment, les croissances de $f(x)$ et de $g(x)$ seront dites *égales;* si ce rapport tend vers zéro, la croissance de $f(x)$ sera dite *inférieure* à la croissance de $g(x)$. Il importe de remarquer qu'étant données deux fonctions positives croissantes, la croissance de l'une d'elles n'est pas nécessairement supérieure, égale ou inférieure à celle de l'autre, car le rapport $\frac{f(x)}{g(x)}$ peut, lorsque x augmente indéfiniment, prendre alternativement des valeurs aussi grandes qu'on veut et aussi petites qu'on veut. C'est là une différence importante entre les croissances et les grandeurs; mais cette difficulté ne se présentant pas avec les fonctions usuelles, on peut la réserver dans une première étude [1] et supposer que l'on considère un ensemble de fonctions croissantes positives, telles que deux quelconques d'entre elles soient *comparables* (c'est-à-dire aient des croissances, ou égales, ou inégales dans un sens bien déterminé). On peut alors se demander s'il est possible de créer pour un tel système une théorie analogue à celle de la mesure des longueurs. Les remarquables travaux de Paul du Bois-Reymond ont montré qu'il n'en est pas ainsi et que l'étude de la croissance des fonctions est beaucoup plus compliquée que l'étude de la mesure des grandeurs continues. Malgré ce résultat négatif, les recherches de Paul du Bois-Reymond n'ont pas été infructueuses; elles l'ont conduit à de nombreux résultats positifs dont le plus important nous paraît être le théorème que nous allons maintenant démontrer [2].

Du Bois-Reymond a obtenu ce théorème en cherchant à construire une suite de fonctions croissantes ayant, par rapport à l'ensemble des autres, les propriétés fondamentales du système des nombres entiers qui servent de base à l'algorithme d'Euclide. D'une manière précise, il a recherché si l'on peut trouver une suite dénombrable de fonctions croissantes

$$(\varphi) \qquad \varphi_1(x), \quad \varphi_2(x), \quad \varphi_3(x), \quad \ldots, \quad \varphi_n(x), \quad \ldots,$$

telles qu'on ait

$$\varphi_1(x) < \varphi_2(x) < \varphi_3(x) < \ldots$$

et que, $\psi(x)$ étant une fonction croissante quelconque donnée, il existe un indice m donnant lieu à l'inégalité

$$\psi(x) < \varphi_m(x).$$

---

(1) On serait amené, pour la résoudre, à reprendre en la complétant la théorie des *enveloppes d'indétermination* de Paul du Bois-Reymond.

(2) M. Pincherle a fait aussi sur la croissance des fonctions des recherches intéressantes; mais nous ne pouvons les exposer sans dépasser les limites de cette Note.

Le théorème de Paul du Bois-Reymond consiste en ce qu'*on ne peut pas trouver de suite telle que* (φ). Sous une forme plus nette, *étant donnée une suite denombrable quelconque* (φ) *de fonctions croissantes, on peut trouver effectivement une fonction croissante* $\psi(x)$, *telle qu'on ait, quel que soit m*,

$$\psi(x) > \varphi_m(x).$$

C'est sous cette forme positive que nous allons démontrer le théorème de Paul du Bois-Reymond; c'est d'ailleurs sous cette forme surtout qu'il peut rendre des services en Analyse. Mais nous avons tenu à indiquer le lien qui le rattachait à la théorie de la mesure des grandeurs, ne serait-ce que pour faire voir, par un exemple précis, comment l'étude approfondie des principes d'une théorie élémentaire, telle que celle de la mesure des grandeurs, peut conduire à des résultats analytiques importants.

Nous considérons donc une suite de fonctions croissantes

$$(\varphi) \qquad \varphi_1(x), \quad \varphi_2(x), \quad \ldots, \quad \varphi_n(x), \quad \varphi_{n+1}(x), \quad \ldots.$$

Par hypothèse, la croissance de $\varphi_{n+1}(x)$ est supérieure à celle de $\varphi_n(x)$, c'est-à-dire que $\frac{\varphi_{n+1}(x)}{\varphi_n(x)}$ augmente indéfiniment avec x; on ne peut en conclure que, pour toute valeur de x, la valeur numérique de $\varphi_{n+1}(x)$ est supérieure à celle de $\varphi_n(x)$. Mais on peut remplacer la suite (φ) par une nouvelle suite de fonctions croissantes

$$(\psi) \qquad \psi_1(x), \quad \psi_2(x), \quad \ldots, \quad \psi_n(x), \quad \ldots,$$

ayant les propriétés suivantes :

1° Deux fonctions de même indice $\varphi_n(x)$ et $\psi_n(x)$ coïncident à partir d'une certaine valeur de x (qui dépendra en général de n);

2° *Pour toute valeur de* x la fonction $\psi_n(x)$ a une valeur supérieure ou égale à la valeur des fonctions $\varphi_n(x)$, $\psi_{n-1}(x)$, $\ldots$, $\psi_1(x)$.

Il suffit de montrer qu'on peut déterminer successivement les fonctions $\psi_n(x)$ de manière qu'elles aient ces propriétés. Supposons que nous ayons pu déterminer les n premières; nous allons voir que nous pourrons sans difficulté déterminer la fonction $\psi_{n+1}(x)$. En effet, par hypothèse, la croissance de $\psi_{n+1}(x)$ est supérieure à celle de $\varphi_n(x)$ et, par suite, à celle de $\psi_n(x)$, qui coïncide avec $\varphi_n(x)$ pour les valeurs suffisamment grandes de x.

Donc, à partir d'une valeur suffisamment grande de x, la valeur de $\varphi_{n+1}(x)$ est supérieure à celle de $\psi_n(x)$. Dès lors, si l'on prend, pour chaque valeur de x, la valeur de $\psi_{n+1}(x)$ égale à la plus grande des valeurs des deux fonctions $\varphi_{n+1}(x)$, $\psi_n(x)$, on sera certain qu'à partir d'une certaine valeur de x la fonction $\psi_{n+1}(x)$ coïncidera avec $\varphi_{n+1}(x)$; elle satisfera d'ailleurs visiblement aux autres conditions exigées.

Considérons maintenant la suite (ψ) et définissons la fonction $\psi(x)$ par les égalités suivantes

$$\psi(n) = \psi_n(n).$$

Ces égalités donnent la valeur de la fonction $\psi(x)$ pour toutes les valeurs entières de x; on peut la définir pour les valeurs non entières par interpolation linéaire (1). Il résulte évidemment des hypothèses faites sur la suite (ψ) que la fonction $\psi(x)$ est croissante; je dis qu'*il en résulte aussi que sa croissance est supérieure à celle de la fonction* $\varphi_m(x)$, *quel que soit l'entier m*. En effet, pour $x > m+1$, la valeur de $\psi(x)$ est visiblement supérieure à celle de $\psi_{m+1}(x)$; donc le rapport $\frac{\psi(x)}{\varphi_m(x)}$ est supérieur au rapport $\frac{\psi_{m+1}(x)}{\varphi_m(x)}$ et ce dernier augmente indéfiniment lorsque x augmente indéfiniment, car, à partir d'une certaine valeur de x, il coïncide avec le rapport $\frac{\varphi_{m+1}(x)}{\varphi_m(x)}$.

Le théorème de Paul du Bois-Reymond est donc démontré.

II. — La formation d'une échelle de types croissants.

Ce théorème nous fournit le moyen, étant donnée une suite dénombrable de fonctions telles que la croissance de chacune d'elles est supérieure à la croissance des précédentes, de trouver une fonction dont la croissance est superieure à celle de chacune des fonctions de la suite.

Nous dirons qu'un ensemble E de fonctions $\varphi(x)$ constitue une *échelle de types croissants* (2) si deux fonctions quelconques de cet ensemble sont comparables entre elles et si, de plus, une fonction croissante *quelconque* $\psi(x)$ étant donnée, il existe dans l'ensemble une fonction $\varphi(x)$ supérieure à $\psi(x)$. La connaissance de l'ensemble E permettrait évidemment de faire un premier pas dans la théorie de la mesure des croissances : le théorème de du Bois-Reymond exprime qu'un tel ensemble ne peut pas être dénombrable. Nous allons néanmoins chercher à en construire un, ne serait-ce que pour donner une idée des difficultés qu'on rencontre lorsqu'on veut définir un ensemble non dénombrable sans faire appel à l'intuition du continu.

Remarquons d'abord que, si l'on a une fonction croissante $\varphi(x)$, dont la croissance est supérieure à celle de x, la fonction $\varphi[\varphi(x)]$ a une croissance supérieure à celle de $\varphi(x)$; car, si l'on pose $\varphi(x) = X$, on a

$$\frac{\varphi[\varphi(x)]}{\varphi(x)} = \frac{\varphi(X)}{X}$$

(1) C'est-à-dire en posant pour $n \leqq x \leqq n+1$

$$\psi(x) = \psi(n) + [\psi(n+1) - \psi(n)](x-n).$$

(2) Deux fonctions dont les croissances seront égales sont dites appartenir au même type.

et, par hypothèse, le dernier rapport augmente indéfiniment avec X ; or X augmente indéfiniment lorsque x augmente indéfiniment.

Si l'on pose

$$\varphi[\varphi(x)] = \varphi_2(x),$$
$$\varphi_{m-1}[\varphi(x)] = \varphi_m(x),$$

la suite $\varphi(x)$, $\varphi_2(x)$, ..., $\varphi_m(x)$, ... est une suite dénombrable de fonctions telles que la croissance de chacune soit supérieure à celle de la précédente. Nous pouvons supposer, pour fixer les idées, $\varphi(x) = e^x$. D'après le théorème de du Bois-Reymond, il existe une fonction $\psi(x)$ dont la croissance dépasse celle de chacune des fonctions $\varphi_m(x)$; nous savons construire une telle fonction et nous pouvons la désigner (¹) par $\varphi_\omega(x)$. Par hypothèse, l'inégalité

$$m > n \tag{1}$$

avait pour conséquence

$$\varphi_m(x) > \varphi_n(x); \tag{2}$$

d'autre part, nous avons, quel que soit m,

$$\varphi_\omega(x) > \varphi_m(x). \tag{2'}$$

Nous pouvons donc convenir d'écrire

$$\omega > m; \tag{1'}$$

cette inégalité (1)', dans laquelle m est un nombre entier arbitraire, n'étant pas autre chose qu'une manière commode d'écrire l'inégalité (2)'.

Nous poserons

$$\varphi_\omega[\varphi_\omega(x)] = \varphi_{2\omega}(x),$$
$$\varphi_{(m-1)\omega}[\varphi_\omega(x)] = \varphi_{m\omega}(x).$$

Il est clair que la suite des fonctions

$$\varphi_\omega(x),\quad \varphi_{2\omega}(x),\quad \ldots,\quad \varphi_{m\omega}(x),\quad \ldots$$

a les mêmes propriétés que la suite

$$\varphi(x),\quad \varphi_2(x),\quad \ldots,\quad \varphi_m(x),\quad \ldots.$$

Nous pouvons, par le théorème de du Bois-Reymond, trouver une fonction

$$\varphi_{\omega\omega}(x) = \varphi_{\omega^2}(x)$$

dont la croissance soit supérieure à celle de chacune des fonctions $\varphi_{m\omega}(x)$.

(¹) En ce moment, nous introduisons cette notation, sans qu'il soit nécessaire d'expliquer quel sens on peut lui attribuer : c'est une simple *notation commode*. Nous verrons plus loin comment elle se rattache à la théorie des nombres transfinis de M. G. Cantor.

En opérant sur $\varphi_{\omega^2}(x)$ de la même manière que sur $\varphi_{\omega}(x)$ on peut obtenir une fonction [1] $\varphi_{\omega^3}(x)$, puis à l'aide de celle-ci une fonction $\varphi_{\omega^4}(x)$, et généralement une fonction $\varphi_{\omega^m}(x)$ dépendant d'un entier m. De l'ensemble de ces fonctions le même théorème permet de déduire une fonction $\varphi_{\omega^\omega}(x)$ dont la croissance dépasse celle de chacune d'elles. On peut continuer ainsi *indéfiniment*, c'est-à-dire, chaque fois que la répétition d'un procédé déterminé aura fourni une infinité dénombrable de fonctions, en obtenir encore une nouvelle par le théorème de du Bois-Reymond. On voit que ce théorème joue à peu près le même rôle que, dans la formation de la suite des nombres entiers, le principe : après chaque nombre entier il y en a un autre.

Mais il importe de remarquer une différence essentielle : d'après sa nature même, le procédé de Paul du Bois-Reymond sera applicable du moment qu'on aura un ensemble dénombrable de fonctions. Or, si nous nous donnons un ou plusieurs procédés *déterminés* pour former des fonctions à croissance de plus en plus rapide, et si nous les appliquons *indéfiniment*, il est clair que nous obtiendrons toujours une infinité dénombrable de fonctions ; car *répéter indéfiniment* une opération ne peut, pour l'instant, être autre chose que la répéter n fois, et faire croître l'entier n indéfiniment. Par suite, si l'on procède ainsi, et si l'on part d'un nombre limité d'opérations (ou même d'une infinité dénombrable), on n'obtiendra jamais qu'une infinité dénombrable de fonctions.

Nous aboutissons ainsi à une antinomie : d'une part, l'application *indéfinie* du théorème de Paul du Bois-Reymond, joint à tels autres procédés d'itération qu'on pourra indiquer, ne peut fournir qu'une infinité dénombrable de types de croissance ; d'autre part, du moment qu'on a seulement une infinité dénombrable de types, le procédé de du Bois-Reymond est encore applicable. Cette antinomie nous impose la *nécessité logique* d'étendre le sens du mot *indéfiniment ;* pour éviter toute confusion, nous éviterons de modifier le sens de ce mot et préférerons introduire le mot nouveau *transfiniment.*

Répéter *transfiniment* l'application du procédé de du Bois-Reymond, ce sera la répéter chaque fois qu'on aura une infinité dénombrable de types croissants, quel que soit le procédé par lequel on a obtenu cette infinité. Par conséquent, par définition même, on obtient ainsi une *infinité non dénombrable* de types ; car, si l'on obtenait seulement une infinité dénombrable, on devrait encore appliquer le même procédé, sans en rester là. L'infinité de types ainsi obtenue sera dite être de la *seconde puissance* et l'on démontre en effet aisément que tout ensemble qui n'est pas de première puissance (p. 13 et 102) a une puissance égale ou supé-

---

[1] A un certain point de vue, il pourrait paraître plus naturel de désigner cette fonction par $\varphi_{\omega^4}(x)$, en regardant ω^4 comme le carré de ω^2, de même que ω^2 est le carré de ω. Mais ici nous sommes libres de choisir arbitrairement nos notations.

rieure à la seconde, ce qui justifie cette dénomination [1]. On peut remarquer que la relation entre la notion de seconde puissance et le théorème de Paul du Bois-Reymond est exactement la même que la relation entre la notion de première puissance et le théorème : après chaque nombre entier, il y en a un autre. Mais l'induction qui nous amène à la notion de la seconde puissance paraît être de nature supérieure à celle qui nous amène à celle de la première puissance, quoique cependant elle s'impose comme tout aussi nécessaire.

Il y a lieu maintenant de se demander si, pour les fonctions croissantes. il existe une proposition analogue à l'axiome d'Archimède que nous avons rappelé au début de cette Note, c'est-à-dire si, *étant donnée une fonction croissante quelconque*, on finira par la dépasser en répétant *transfiniment* les procédés que nous avons indiqués. Il nous semble que c'est là un *axiome* qui doit être admis au même titre que l'axiome d'Archimède et d'une manière tout à fait générale; il est, en tout cas, certain que cette proposition n'est pas douteuse si, par les mots : *fonction quelconque*, on entend une fonction qui puisse être *effectivement définie*. c'est-à-dire telle qu'on puisse, par un nombre limité d'opérations, calculer. avec une approximation donnée, sa valeur pour une valeur donnée de la variable.

Désignons par (S) la suite transfinie de fonctions croissantes que nous venons de définir. Cette suite a les propriétés fondamentales suivantes :

1° Deux fonctions quelconques sont comparables entre elles ;

2° Étant donnée une fonction quelconque de S, il y en a une qui la suit immédiatement ;

3° Une fonction quelconque étant donnée, il y a dans S une infinité de fonctions qui lui sont supérieures.

Cette suite S possède ainsi quelques-unes des propriétés fondamentales de la suite des nombres entiers. On peut en déduire un ensemble Σ qui possédera de même quelques-unes des propriétés fondamentales de l'ensemble des nombres rationnels. Il suffit pour cela de procéder exactement de même que pour obtenir ce dernier ensemble : la considération des fonctions *inverses* des fonctions de S donnera des fonctions croissant de moins en moins vite; en les multipliant par la variable x, on aura des fonctions à croissance *plus rapide* que x, mais aussi peu que possible; par le procédé de l'itération, répété transfiniment, on formera un ensemble Σ de fonctions à croissance plus rapide que x et moins rapide que celle de x^2,

---

[1] En effet, tout ensemble non dénombrable a évidemment la propriété suivante : Si l'on en extrait un ensemble dénombrable quelconque, il y subsiste encore un élément, et l'on peut répéter cette opération *transfiniment*, car, si l'on ne la pouvait répéter qu'*indéfiniment*, l'ensemble serait dénombrable. Cela suffit pour que nous soyons assurés que l'ensemble non dénombrable a une puissance égale ou supérieure à celle que nous appelons la *seconde*.

par exemple. Cet ensemble Σ sera d'ailleurs de seconde puissance et aura les propriétés suivantes :

1°. Deux fonctions quelconques de Σ' seront comparables entre elles.

2° Si l'on considère une fonction quelconque, comparable à toutes les fonctions de Σ, et dont la croissance est comprise entre celle de x et celle de x^2, la croissance de cette fonction sera entièrement définie par la suite transfinie des fonctions de Σ dont la croissance est plus grande et la suite transfinie des fonctions de Σ dont la croissance est plus petite.

D'ailleurs, réciproquement, tout mode de division des fonctions de Σ en deux ensembles, tels que chaque fonction du premier ait une croissance inférieure à toute fonction du second définit *un mode de croissance*, mais à ce mode ne correspond pas nécessairement une *fonction*. De même que toute division en deux classes de l'ensemble des nombres rationnels définit une *grandeur*, mais à cette grandeur ne correspond pas nécessairement un *nombre*, tant qu'on n'a pas *convenu* d'appeler *nombre* les nombres incommensurables. Il y aurait lieu de même ici, pour compléter ce *continu fonctionnel*, analogue au *continu linéaire*, d'introduire des *fonctions idéales*, analogues aux *nombres incommensurables*.

Une *fonction idéale*, c'est *un mode de division de l'ensemble* Σ *en deux classes telles que toute fonction de la première classe soit inférieure à toute fonction de la seconde classe et telles, de plus, qu'il n'y ait pas dans la première classe de fonction supérieure à toutes les autres, ni dans la seconde classe de fonction inférieure à toutes les autres.*

Par exemple, rangeons dans la première classe les fonctions $\varphi(x)$ telles que l'intégrale $\int_0^\infty \frac{dx}{\varphi(x)}$ n'ait pas de sens et dans la seconde les fonctions telles que l'intégrale $\int_0^\infty \frac{dx}{\varphi(x)}$ ait un sens. Nous aurons ainsi défini une *fonction idéale* telle que, si on la désigne par $\theta(x)$, l'intégrale

$$\int_0^\infty \frac{dx}{\theta(x)}$$

est a la fois pourvue et dépourvue de sens. Cette propriété est tout aussi absurde pour celui qui regarderait $\theta(x)$ comme *une véritable fonction*, que la suivante, pour celui qui ne considérerait que de *véritables nombres :* il existe un nombre dont le carré est égal à 2.

D'ailleurs, il ne faut pas oublier que les fonctions de Σ sont supposées toutes comparables entre elles. Si nous sommes partis, pour la formation le S, de la fonction exponentielle e^x, nous dirons que les fonctions de S et de Σ appartiennent au type exponentiel et les fonctions croissantes qui s'introduisent naturellement leur seront comparables.

Nous avons dit que l'ensemble Σ est de la seconde puissance, c'est-

à-dire d'une puissance *inférieure, ou égale*, à celle du continu; il est aisé de voir que si on le complète à l'aide des *fonctions idéales*, on obtient un ensemble dont la puissance est *égale ou supérieure* à celle du continu. Nous nous contenterons de ces indications ([1]).

Que l'ensemble de toutes les fonctions croissantes (non idéales) ait une puissance égale à celle du continu, c'est une conséquence du fait que l'ensemble des fonctions continues a la puissance du continu, et il est clair que nous pouvons nous borner, pour définir la croissance, à considérer les fonctions continues croissantes.

Donc, la puissance de l'ensemble des fonctions croissantes est au plus égale à celle du continu; mais, d'autre part, elle lui est au moins égale, car cet ensemble renferme en particulier l'ensemble des fonctions x^α, où α est un nombre incommensurable quelconque : donc elle lui est égale. Ce point nous a paru utile à établir en toute rigueur, car il avait été contesté par du Bois-Reymond dans sa *Théorie générale des fonctions*.

III. — Les nombres de M. G. Cantor.

Nous sommes maintenant à même d'exposer, comme nous la comprenons, la formation des nombres *plus grands que l'infini* introduits par M. G. Cantor. Citons d'abord textuellement quelques passages de M. G. Cantor (*Acta mathematica*, t. II, p. 385-399) :

« ..... La série (I) des nombres entiers réels et positifs $1, 2, 3, \ldots, \nu, \ldots$ doit sa formation à la répétition et à la réunion d'unités qu'on a prises pour point de départ et que l'on considère comme égales..... La formation des nombres entiers réels *finis* repose donc sur le principe de l'addition d'une unité à un nombre *déjà formé;* j'appelle *premier principe* de formation ce moment qui, comme nous le verrons bientôt, joue aussi un rôle essentiel dans la production des nombres entiers supérieurs. Le nombre des nombres ν de la classe (I), formés de cette manière, est infini et parmi tous ces nombres il n'y en a pas qui soit plus grand que tous les autres. Il serait donc contradictoire de parler d'un nombre maximum de la classe (I): toutefois, on peut d'autre part imaginer un nouveau nombre, que nous appellerons ω, et qui *servira à exprimer que l'ensemble* (I) *est donné d'après la loi dans sa succession naturelle*. On peut même se représenter le nouveau nombre ω comme la limite vers laquelle tendent les nombres ν, à condition d'entendre par là que ω sera le *premier* nombre entier qui *suivra tous* les nombres ν, en sorte qu'il faut le déclarer supérieur à *tous* les nombres ν. En associant le nombre ω avec les unités

---

([1]) La première considération des fonctions idéales est due à Paul du Bois-Reymond: j'ai cherché à compléter un peu ses trop brèves indications. Dans un Mémoire récent, M. Pringsheim a fait, aux idées de Paul du Bois-Reymond sur ce sujet, des objections que je n'ai pu arriver à comprendre.

primitives, on obtient à l'aide du *premier principe* de formation ces nombres plus étendus

$$\omega+1,\quad \omega+2,\quad \ldots,\quad \omega+\nu,\quad \ldots;$$

comme par là on n'arrive encore une fois à aucun nombre maximum, on en imagine un nouveau qu'on peut appeler 2ω et qui sera le *premier* après tous les nombres obtenus jusqu'à présent ν et $\omega+\nu$, ...

» ..... La fonction logique qui nous a donné ces nombres ω et 2ω est évidemment différente du premier principe de formation; je l'appelle *deuxième* principe de formation des nombres réels entiers et je définis mieux ce principe en disant : *Étant donnée une succession quelconque déterminée de nombres entiers réels définis, parmi lesquels il n'y en a pas qui soit plus grand que tous les autres, on pose, en s'appuyant sur ce deuxième principe de transformation, un nouveau nombre qu'on regarde comme la limite des premiers, c'est-à-dire qui est défini comme étant immédiatement supérieur à tous ces nombres.....*

» ..... La formation de nouveaux nombres, comme on le voit, est *sans fin;* en suivant les deux principes de formation on obtient toujours de nouveaux nombres et de nouvelles séries de nombres, avec une succession *parfaitement déterminée.*

» *On pourrait donc croire* d'abord que nous allons nous perdre à l'indéfini dans cette formation de nouveaux nombres entiers infinis déterminés et que nous ne sommes pas en état d'*arrêter provisoirement* ce procédé sans fin pour arriver par là à une *limitation semblable à celle que nous avons trouvée, en fait, dans un certain sens, par rapport à l'ancienne classe de nombre* (I); là, on n'employait que le premier principe de formation et l'on ne pouvait pas sortir de la série (I). Mais le deuxième principe de formation ne devait pas seulement nous conduire au delà du système de nombres employé jusqu'à présent; il nous apparaît encore certainement comme un moyen qu'on peut combiner avec le premier principe de formation pour arriver à *pouvoir franchir toute limite* dans la formation abstraite des nombres réels entiers.

» *Mais si nous remarquons maintenant que tous les nombres obtenus jusqu'à présent et ceux qui les suivent immédiatement remplissent une certaine condition*, nous verrons que cette condition, *si on la pose comme obligatoire pour tous les nombres à former immédiatement*, nous apparaît comme un *troisième* principe qui vient s'ajouter aux deux premiers et que j'appelle *principe d'arrêt et de limitation*..... Cette condition est : que *le système des nombres qui se trouvent dans la suite des nombres avant celui que l'on considère* et à partir de 1 soit de la même puissance que la première classe de nombres (I)... . La deuxième classe de nombres (II) définie par l'adjonction de ce principe n'acquiert pas seulement une puissance plus élevée que (2), mais *précisément la puissance immédiatement supérieure* et, par conséquent, la *deuxième puissance.* »

On voit l'analogie qu'il y a entre le *deuxième principe* de M. G. Cantor et le *théorème* de Paul du Bois-Reymond et, en même temps, les différences profondes qui les séparent. Le *principe* est posé *a priori*; par suite, si on l'admet, on doit lui accorder une valeur absolue; il débordera tout, si on ne l'arrête, ce qui est le rôle du *troisième principe*. Le *théorème*, au contraire, n'est pas un postulat; c'est un *fait mathématique* qui ne repose sur aucune considération *a priori*, mais sa puissance est bien plus limitée : il porte en lui-même le principe d'arrêt, car il n'est applicable qu'autant que l'ensemble déjà obtenu est dénombrable. D'autre part, comme son application *indéfinie* conduit seulement à un ensemble dénombrable, auquel il est *encore* applicable, il y a là une antinomie qui, comme nous l'avons vu, ne peut se résoudre qu'en attribuant un sens au mot *transfiniment* et admettant, par suite, qu'en appliquant *transfiniment* le théorème on aura un ensemble de fonctions de puissance supérieure à la première obtenu ainsi d'une manière nécessaire (1).

On peut se demander si, au contraire, il n'y a pas beaucoup d'illusion dans l'idée que nous nous sommes faite, après M. G. Cantor, de la puissance du deuxième principe de formation.

Il semble en effet, quand on y regarde d'un peu plus près, que ce principe n'ait nullement le pouvoir de créer par lui-même de nouvelles *puissances*, c'est-à-dire de nous élever, par sa seule force, au delà de la première puissance. Il est bien clair que, si nous utilisons le premier et le deuxième principe, l'un après l'autre, une infinité dénombrable de fois, nous n'obtiendrons jamais qu'un ensemble de première puissance.

D'une manière plus générale, quel que soit le procédé *déterminé* de formation qu'on indique, du moment que ce procédé sera *complètement exprimable* au moyen d'un nombre fini de mots, parmi lesquels peut

---

(1) Une difficulté assez sérieuse provient de ce que le théorème de Paul du Bois-Reymond ne conduit pas à une fonction déterminée, car il peut être appliqué d'une infinité de manières. Il ne conduit surtout pas à la *plus petite* des fonctions supérieures aux fonctions données. Mais on peut remarquer que, pour former la fonction $\psi(x)$ au moyen de la suite $\psi_1(x), \ldots, \psi_n(x), \ldots$ (en reprenant les notations de la page 113), on se sert en réalité d'une certaine fonction croissante; nous avons posé

$$\psi(n) = \psi_n(n);$$

nous aurions pu prendre, plus généralement,

$$\psi[\theta(n)] = \psi_n[\theta(n)]$$

$\theta(n)$ étant une fonction croissante quelconque. Mais il est naturel de ne prendre pour $\theta(n)$ qu'une des fonctions croissantes qui ont été définies jusque-là, et non pas une fonction à croissance tellement rapide qu'on ne l'ait pas encore atteinte. On peut prendre aussi pour $\theta(n)$ la fonction inverse des fonctions croissantes déjà considérées; mais on n'a le choix qu'entre une infinité dénombrable, ce qui est l'essentiel.

figurer le mot *indéfiniment*, on n'obtiendra jamais qu'un ensemble de première puissance, si le mot *indéfiniment* signifie toujours : *aussi souvent qu'il y a de nombres entiers*. Et nous n'avons pas le droit de donner un autre sens au mot *indéfiniment*, si nous n'avons pas la notion d'une puissance supérieure à la première; or c'est précisément cette notion qu'il s'agit d'acquérir : nous tournons dans un cercle.

Si nous introduisons l'expression *transfiniment*, en lui donnant le sens précis de *aussi souvent qu'il y a d'éléments dans un ensemble de deuxième puissance*, nous pourrons dire que l'application transfinie du deuxième principe donnera des nombres de deuxième classe, mais on n'obtiendra jamais ainsi un ensemble de nombres de *troisième* puissance. Il faudrait avoir déjà l'idée d'un ensemble ayant cette puissance, afin de répéter l'application du deuxième principe aussi souvent qu'il y a d'éléments dans cet ensemble.

En d'autres termes, le deuxième principe de formation ne peut nous faire acquérir la notion d'une *puissance* que nous n'aurions pas déjà; et il semble bien douteux que nous ayons une idée quelque peu précise de ce que peut être une puissance qui dépasse la deuxième ([1]). Il est bien probable que, avant d'acquérir une telle idée, si nous y arrivons jamais, il aura été nécessaire d'étudier d'une manière approfondie les ensembles de deuxième puissance, de telle manière que nous sachions en user aussi commodément que des ensembles dénombrables. Le jour ou ce résultat, dont l'importance pour la Science serait considérable, sera atteint, on aura une base solide pour l'étude des puissances supérieures; mais on ne peut nier que, actuellement, l'expression *transfiniment* n'ait encore pour nous un sens moins précis que l'expression *indéfiniment*, de sorte que nos connaissances précises sur les puissances diverses n'excèdent guère la remarque suivante : il y a des ensembles dénombrables et des ensembles non dénombrables, cette dernière notion étant surtout négative.

([1]) Je laisse de côté ici, sans m'inquiéter de savoir si elles sont ou non distinctes des puissances des classes de nombres, la puissance du continu et les puissances qui s'y rattachent (ensembles de fonctions, etc.) dont nous acquerrons la notion par des procédés tout différents. Il se présente, en effet, lorsqu'on emploie ces procédés pour définir les puissances, une difficulté très grave : on définit des puissances de plus en plus grandes, mais on ne sait pas si elles sont *consecutives*, c'est-à-dire s'il n'y a pas, par exemple, une puissance intermédiaire entre celle du continu et celle de l'ensemble des fonctions discontinues.

NOTE III.

LA NOTION DE FONCTION EN GÉNÉRAL.

La notion de fonction est extrêmement générale et comprend plusieurs notions particulières qu'on peut considérer séparément : c'est ainsi qu'on peut distinguer, entre les fonctions d'une ou de plusieurs variables, les fonctions continues et discontinues, les fonctions analytiques d'une variable complexe et les fonctions de variable réelle, les fonctions ayant ou n'ayant pas de dérivée, etc. Nous nous proposons de jeter un coup d'œil d'ensemble sur ces diverses notions. et, en particulier, de tâcher de nous rendre compte en quoi elles diffèrent et en quoi elles sont analogues.

I. — Les fonctions discontinues.

Occupons-nous d'abord de la notion la plus générale de la fonction, sans aucune restriction : ce sera une fonction discontinue. Une fonction f de trois variables x, y, z, par exemple, sera définie par cette *seule* condition, qu'à tout système [1] de valeurs des trois variables x, y, z correspond une valeur bien déterminée [2] de $f(x, y, z)$. Il est aisé de voir que le nombre des variables importe peu dans une telle notion; nous savons, en effet, que l'ensemble des points de l'espace a même puissance que l'ensemble des points d'une droite compris entre 0 et 1. Si, au système de valeurs x, y, z correspond ainsi la valeur X comprise entre 0 et 1, il suffit de poser

$$f(x,y,z) = \varphi(X),$$

pour définir une fonction $\varphi(X)$ d'une seule variable X, qui correspond d'une manière bien déterminée à la fonction $f(x,y,z)$. Cette correspondance est d'ailleurs réciproque et univoque : deux fonctions étant regardées

(1) On pourrait considérer seulement les systèmes de valeurs appartenant à un certain domaine; pour les autres, la fonction ne serait pas définie.

(2) On pourrait supposer, plus généralement, qu'il y a plusieurs valeurs, ou même une infinité dénombrable; on verra aisément que les conséquences qu suivent n'en seraient pas modifiées.

comme distinctes lorsqu'elles ont une valeur différente en l'*un* au moins des points où elles sont définies, à deux fonctions distinctes $f(x, y, z)$ correspondent deux fonctions distinctes $\varphi(X)$ et réciproquement.

Par conséquent, *l'ensemble de toutes les fonctions discontinues d'un nombre quelconque de variables a même puissance que l'ensemble des fonctions discontinues d'une variable réelle* X *assujettie à rester comprise entre* 0 *et* 1.

Il est facile de ramener la puissance de ce dernier ensemble à une puissance définie d'une manière un peu plus simple. Remarquons d'abord qu'on n'introduit pas de restriction sérieuse en supposant que la fonction $\varphi(X)$ a constamment une valeur comprise ([1]) entre 0 et 1. Si nous écrivons cette valeur dans le système de numération de base 2, elle sera de la forme

$$(\alpha) \qquad \varphi(X) = 0{,}10100111\ldots,$$

tous les chiffres décimaux étant égaux à 0 ou à 1.

Remarquons que, X étant un nombre quelconque compris entre 0 et 1, tous les nombres positifs sont de la forme $n + X$, n étant un nombre entier. Cela étant, nous pouvons définir un ensemble de points de la manière suivante : *Le point* $n + X$ *appartiendra à l'ensemble* si, dans la valeur (α) de $\varphi(X)$, le $n^{\text{ième}}$ chiffre décimal est 1; il ne lui appartiendra pas si ce $n^{\text{ième}}$ chiffre décimal est 0. Il est clair que chaque fonction $\varphi(X)$ permet de définir un tel ensemble; d'ailleurs, cet ensemble est unique si la fonction ne prend aucune valeur commensurable de la forme $\frac{p}{2^m}$, car un tel nombre peut se mettre de deux manières différentes sous la forme (α), en vertu de l'identité

$$1 = \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} + \ldots = 0{,}1111\ldots.$$

Réciproquement à un tel ensemble correspond une fonction unique $\varphi(X)$;

([1]) En effet, il suffit de remplacer $\varphi(X)$ par la fonction $\psi(X)$ ainsi définie : Si

$$\begin{aligned}
0 < \varphi(X) < 1, &\qquad \psi(X) = \frac{1}{4}\varphi(X),\\
-1 < \varphi(X) < 0, &\qquad \varphi(X) = \frac{1}{4}\varphi(X) + \frac{1}{2},\\
1 < \varphi(X) < +\infty, &\qquad \varphi(X) = \frac{1}{2} + \frac{1}{4\varphi(X)},\\
-\infty < \varphi(X) < -1, &\qquad \psi(X) = 1 + \frac{1}{4\varphi(X)},
\end{aligned}$$

voit aisément que, réciproquement, la connaissance de la valeur de $\psi(X)$ ermine celle de $\varphi(X)$.

donc, *l'ensemble des fonctions* $\varphi(X)$ *a une puissance au plus égale à la puissance de l'ensemble des ensembles de points réels positifs*. Mais il est clair que cette puissance est la même que l'ensemble des ensembles de points compris entre 0 et 1 ; et, comme un tel ensemble peut être considéré comme correspondant à une fonction $\varphi(X)$ prenant seulement les deux valeurs 0 et 1, on voit que l'ensemble de ces ensembles a une puissance inférieure ou égale à l'ensemble des fonctions $\varphi(X)$; ces deux puissances sont donc égales.

Nous voyons donc, en définitive, que (¹) *l'ensemble des fonctions discontinues d'un nombre quelconque de variables réelles a même puissance que l'ensemble des fonctions discontinues d'une variable réelle, comprise entre* 0 *et* 1, *ces fonctions prenant seulement la valeur* 0 *ou la valeur* 1, *ou*, ce qui revient au même, *que l'ensemble des ensembles de points compris entre* 0 *et* 1.

Nous avons déjà montré que cette puissance est supérieure à la puissance du continu; mais il ne nous paraît pas inutile d'insister encore sur la difficulté qu'il y a à définir un tel ensemble. Étant données, en effet, deux fonctions discontinues $\varphi(X)$ et $\varphi_1(X)$ que nous pouvons, d'après ce qui précède, supposer égales toujours à 0 ou à 1, le problème de reconnaître si elles sont identiques ou différentes présente une difficulté toute particulière. Il n'est pas possible, en effet, de fixer une méthode telle que, si elles sont différentes, on en soit assuré *sûrement* au bout d'un nombre *fini* d'opérations; cela tient à ce que le continu n'est pas dénombrable. Si les fonctions $\varphi(X)$ et $\varphi_1(X)$ étaient continues, il en serait tout autrement; en effet, si deux telles fonctions coïncident pour les valeurs *commensurables* de X, elles coïncident pour toute valeur de X; si donc elles ne sont pas identiques, on s'en apercevra sûrement par la seule considération des valeurs commensurables : *il existe* un nombre commensurable $\frac{p}{q}$ tel que $\varphi\left(\frac{p}{q}\right)$ et $\varphi_1\left(\frac{p}{q}\right)$ diffèrent, c'est-à-dire aient seulement m chiffres décimaux communs, m étant un nombre déterminé.

Dès lors, si l'on range tous les nombres commensurables en une suite

$$u_1, u_2, \ldots, u_n, \ldots,$$

ce qu'on sait faire, et si l'on calcule les valeurs $\varphi(u_1), \ldots, \varphi(u_n)$, $\varphi_1(u_1), \ldots, \varphi_1(u_n)$, avec n chiffres décimaux exacts, il arrivera nécessai-

(¹) Si, employant un langage introduit par M. G. Cantor, nous désignons par a le *nombre cardinal* des points compris entre 0 et 1, ce théorème peut être traduit par l'égalité

$$a^{a^n} = 2^a,$$

n étant un nombre entier quelconque; on pourrait d'ailleurs supposer aussi que n est un nombre de la deuxième classe.

rement, si l'on fait successivement cette opération pour

$$n = 1, 2, 3, \ldots,$$

que, pour une valeur *finie* de n, c'est-à-dire au bout d'un nombre *fini* d'opérations, on sera assuré que les deux fonctions ne sont pas identiques.

D'ailleurs, il est évident que ce nombre ne sera pas connu d'avance et que, par suite, on ne pourra pas, de calculs aussi longs qu'ils soient, conclure à l'identité des deux fonctions. Mais, si elles diffèrent, on est sûr de s'en apercevoir avec assez de persévérance. Il n'y a rien de pareil pour les fonctions discontinues : une telle fonction est définie *par une infinité non dénombrable de conditions;* en pratique, cela revient à dire qu'il est *impossible* de la *définir*. Il y aurait lieu de distinguer parmi les fonctions discontinues les plus générales, qui paraissent devoir être exclues, au moins pour l'instant, des considérations mathématiques, des fonctions dont la discontinuité est assujettie à des restrictions. Ces *restrictions doivent être de nature telle que la fonction puisse être entièrement définie par une infinité dénombrable de conditions*. L'ensemble des fonctions satisfaisant à ces conditions restrictives a alors même puissance que le continu. Parmi les fonctions satisfaisant à des conditions restrictives de ce genre, je citerai les fonctions discontinues seulement en une infinité dénombrable de points, que plusieurs géomètres ont considérées avec profit. Mais je ne puis m'étendre longuement sur ce sujet; je tenais seulement à indiquer la restriction indispensable qu'on doit apporter à l'idée générale de fonction, si l'on veut pouvoir l'utiliser ([1]).

II. — Les fonctions définies par des conditions dénombrables.

On peut évidemment imaginer une infinité de classes de fonctions, telles que les fonctions de chaque classe soient entièrement définies à l'aide d'une infinité dénombrable de conditions ([2]) de sorte que la puissance de l'ensemble formé par les fonctions d'une classe est précisément la puissance du continu. Nous nous proposons d'étudier quelques-unes de ces classes de fonctions, en particulier au point de vue des rapports qu'elles ont entre elles : on apercevra aisément, d'ailleurs, ce qui, dans

---

([1]) Cela ne veut pas dire qu'on ne puisse trouver quelques résultats s'appliquant à *toutes* les fonctions et pouvant rendre des services; tel est, par exemple, le beau théorème de M. Darboux sur l'intégrale définie : *Il existe toujours* une intégrale par excès et une intégrale par défaut (*voir* JORDAN, *Cours d'Analyse*). Mais de tels résultats sont forcément rares et ce théorème lui-même, malgré sa beauté théorique, n'a d'applications possibles que si l'on restreint beaucoup la généralité de la notion de fonction.

([2]) Bien entendu ces conditions sont *simples;* on peut supposer, si l'on veut, que chaque condition se traduit par le fait qu'on donne un nombre entier.

nos considérations s'appliquerait sans modification à telles autres classes qu'on voudrait considérer.

On appelle *fonction analytique* de plusieurs variables x, y, z toute fonction qui peut être représentée par un développement de la forme

$$(1)\qquad \mathfrak{P}(x-a, y-b, z-c) = \Sigma_{\alpha,\beta,\gamma} A_{\alpha,\beta,\gamma}(x-a)^{\alpha}(y-b)^{\beta}(z-c)^{\gamma}$$

ordonné suivant les puissances positives de $x-a$, $y-b$, $z-c$, et convergent dans certains cercles, c'est-à-dire sous les conditions

$$|x-a| < r;\quad |y-b| < r',\quad |z-c| < r'',$$

r, r', r'' étant des nombres positifs déterminés.

La condition que *le développement* (1) *ne diverge pas pour tout système de valeurs* [1] *de x, y, z* est essentielle; elle s'exprime par des inégalités assez compliquées auxquelles doivent satisfaire [2] les coefficients $A_{\alpha,\beta,\gamma}$. Désignons par E_3 l'ensemble des développements tels que (1), les coefficients $A_{\alpha,\beta,\gamma}$ étant absolument quelconques, et par F_3 l'ensemble des développements qui satisfont à la condition dont nous venons de parler : l'ensemble F_3 est précisément l'ensemble des fonctions analytiques à *trois* variables; il est une partie aliquote de l'ensemble E_3.

D'une part il est aisé de montrer que l'ensemble E_3 a la puissance du continu (p. 19); d'autre part, on voit aisément que l'ensemble F a une puissance supérieure ou égale à celle du continu (car, l'un des coefficients, par exemple $A_{0,0,0}$ peut évidemment prendre une valeur arbitraire); donc l'ensemble F_3 a précisément la puissance du continu : mais l'*établissement effectif* d'une correspondance univoque et réciproque, entre les éléments de F_3 et les nombres compris entre 0 et 1, par exemple, paraît un problème ardu [3]; tandis qu'une telle correspondance est, au contraire, très aisée à établir, si, au lieu de F_3, on considère E_3.

Un phénomène analogue se produit lorsque l'on compare les ensembles de fonctions analytiques dépendant d'un nombre différent de variables; nous allons l'étudier avec un peu plus de détail.

Pour simplifier l'écriture, nous allons comparer l'ensemble F_2 des fonctions analytiques à *deux* variables et l'ensemble F_1 des fonctions analytiques à *une* variable. Nous considérerons en même temps les

---

[1] Les systèmes de valeurs, pour lesquels *l'un* au moins des binomes $x-a$ $y-b$, $z-c$ serait nul, sont ici exclus.

[2] On peut consulter à ce sujet une Note de M. E. LEMAIRE, *Bulletin des Sciences mathématiques*, 1897.

[3] Il semble que la seule voie à suivre soit précisément celle qui a servi à démontrer le théorème sur lequel nous nous sommes appuyés (Note I, p. 106), mais cette marche, théoriquement simple, n'est pas sans présenter des difficultés pratiques assez sérieuses.

ensembles E_2 et E_1 formés par les développements non assujettis à la condition de convergence [1]:

$$(E_2) \qquad \Sigma A_{\alpha,\beta} x^\alpha y^\beta,$$

$$(E_1) \qquad \Sigma B_n z^n.$$

Les ensembles F_2 et F_1 sont respectivement des parties aliquotes de E_2 et de E_1; tous ces ensembles ont d'ailleurs la puissance du continu et par suite deux quelconques d'entre eux ont même puissance. Mais nous allons voir qu'il est très aisé d'établir une correspondance univoque et réciproque entre les éléments de E_2 et de E_1, tandis qu'il n'en est pas de même si l'on considère F_2 et F_1, bien qu'on soit assuré que ces ensembles ont même puissance.

Pour établir la correspondance entre les éléments de E_2 et ceux de E_1, il suffit (*cf.* p. 8) de poser

$$A_{\alpha,\beta} = B_n$$

sous la condition

$$(a) \qquad n = \frac{(\alpha+\beta)(\alpha+\beta+1)}{2} + \alpha.$$

Lorsqu'on donne α et β, cette relation permet de calculer n; si l'on donne n, on prendra d'abord pour $\alpha+\beta$ le plus grand nombre entier u vérifiant l'inégalité

$$\frac{u(u+1)}{2} \leqq n,$$

ce qui revient à dire que, v étant la racine carrée à une unité près par défaut de $8n+1$, $2u+1$ est égal à v si v est impair et à $v-1$ si v est pair.

Connaissant $u = \alpha+\beta$, la relation (a) donnera

$$\alpha = n - \frac{u(u+1)}{2}$$

et l'on aura ensuite $\beta = u - \alpha$.

La correspondance entre les éléments de E_2 et ceux de E_1 est ainsi très simplement définie; mais il est aisé de se rendre compte que cette correspondance transforme certaines séries E_2 *divergentes pour toutes les valeurs de x et de y* en séries E_1 *convergentes pour certaines valeurs*

---

[1] Nous supposons que les centres des cercles de convergence coïncident toujours avec l'origine; il n'y aurait pas plus de difficultés à étudier les ensembles plus étendus (mais de même puissance) qu'on obtient en les supposant arbitraires.

de z. Il suffit, en effet, de poser, par exemple,

$$A_{\alpha,\beta} = M_{\alpha,\beta}(\log u)^u,$$

$M_{\alpha,\beta}$ étant un nombre compris entre 1 et 2, par exemple. En effet, si la série E_2 était convergente pour $x = x_0$, $y = y_0$, la série

$$\sum_{\alpha,\beta} |A_{\alpha,\beta}| r^{\alpha+\beta}$$

serait convergente sous la seule condition que le nombre réel r soit inférieur à $|x_0|$ et à $|y_0|$; or, en désignant par M_u la somme des valeurs de $M_{\alpha,\beta}$ pour $\alpha + \beta = u$, cette série peut s'écrire

$$\sum_{u=1}^{u=\infty} M_u (\log u)^u r^u,$$

et il est manifeste qu'elle n'est convergente pour aucune valeur de r, les nombres M_u étant tous supérieurs à *un*.

D'autre part, si l'on pose

$$B_n = M_{\alpha,\beta}(\log u)^u,$$

comme l'on a, par hypothèse,

$$M_{\alpha,\beta} < 2,$$

et, d'après (a),

$$u < 2\sqrt{n},$$

on en conclut

$$B_n < (\log n + 1)^{2\sqrt{n}},$$

et il en résulte aisément que la série

$$\Sigma B_n z^n,$$

est convergente pour $|z| < 1$.

On verrait aisément que toute correspondance simple établie entre les éléments de E_2 et ceux de E_1 donne lieu aux mêmes difficultés quand on considère F_2 et F_1; la correspondance est alors en défaut pour certains éléments de l'un de ces deux ensembles; à chacun de ces éléments ne correspond aucun élément de l'autre ensemble.

Malgré ce grave inconvénient, la correspondance que nous venons de réaliser est intéressante à cause de sa simplicité, et aussi de la propriété suivante, qui est évidente et qu'elle n'est pas d'ailleurs la seule à posséder : *à la somme de deux séries elle fait correspondre la somme des deux séries correspondantes* (E_1). Il est donc naturel de se demander si l'on ne pourrait pas s'en servir utilement, en se résignant à faire abstraction des éléments de F_1 auxquels ne correspond aucun élément de F_2. Mais

nous allons rencontrer une autre difficulté ; nous venons de remarquer que si, aux fonctions $\varphi_1(x, y)$ et $\varphi_2(x, y)$ correspondent respectivement les fonctions $\psi_1(z)$ et $\psi_2(z)$, à la fonction $\varphi_1(x, y) + \varphi_2(x, y)$ correspondra la fonction $\psi_1(z) + \varepsilon\psi_2(z)$. Par suite, à la fonction

$$\varphi(x, y) = \varphi_1(x, y) + \varepsilon\,\varphi_2(x, y),$$

correspondra la fonction

$$\psi(z) = \psi_1(z) + \varepsilon\,\psi_2(z).$$

Si nous supposons que ε soit un infiniment petit, nous pourrons être tentés de dire que la fonction $\varphi(x, y)$ est infiniment voisine de la fonction $\varphi_1(x, y)$ et, par suite, d'énoncer le résultat suivant : *A deux fonctions infiniment voisines $\varphi_1(x, y)$ et $\varphi(x, y)$ correspondent deux fonctions infiniment voisines $\psi_1(z)$ et $\psi(z)$.* Mais, en y regardant de plus près, on se rendra compte que la définition que nous venons de donner implicitement est tout à fait arbitraire et que rien ne la justifie. Il est certain que, suivant le genre de recherches qu'on fera, il peut être commode d'adopter telle ou telle définition des mots : *fonctions infiniment voisines.* Ce n'est point ici le lieu de discuter les diverses définitions qu'on pourrait proposer. Mais il nous semble que, si l'on ne se borne pas à considérer la fonction au voisinage d'un point, il n'y a pas lieu de considérer le développement

$$1 + z + z^2 + \ldots + \varepsilon(1 + 2z + 4z^2 + 8z^3 + \ldots)$$

comme infiniment voisin de celui qu'on obtient en faisant $\varepsilon = 0$; car, quelque petit que soit ε, dans le coefficient de $z^n(1 + \varepsilon 2^n)$, c'est le second terme qui est le plus grand à partir d'une certaine valeur de n.

Une observation analogue s'appliquerait au développement en série de la fonction

$$e^z + \varepsilon e^{z^2}$$

Sans insister sur ces questions, on pressent la nature des difficultés qui se présenteront si l'on tient qu'à deux fonctions infiniment voisines correspondent deux fonctions infiniment voisines. Il faudra choisir les définitions d'une manière convenable, en ayant soin cependant de n'y point mettre trop d'arbitraire. On y arriverait sans doute ici sans trop de peine ; mais avec d'autres classes de fonctions il pourrait en être tout autrement ; la première difficulté que nous avons signalée se présenterait d'ailleurs aussi (1).

Parmi les classes de fonctions qu'on pourrait ainsi considérer, on peut citer l'ensemble des fonctions d'une variable réelle développables en série

---

(1) *Cf.* un intéressant Mémoire de M. Pincherle : *Sur le calcul fonctionnel distributif* (*Math. Annalen*, t. XLIX).

de Fourier :

$$\text{(F)} \qquad a_0 + \Sigma(a_n \cos nx + b_n \sin nx);$$

on peut citer aussi l'ensemble des fonctions continues d'une ou plusieurs variables réelles. Pour donner ces dernières fonctions, il suffit de donner leur valeur pour les systèmes de valeurs toutes commensurables des variables; ou, plus généralement, pour les points d'un ensemble dénombrable dense dans tout le domaine où la fonction doit être définie. Mais ces valeurs ne sont pas arbitraires : le fait que la fonction est continue les assujettit à des conditions d'une nature très compliquée.

Aussi l'établissement effectif d'une correspondance univoque et réciproque entre les éléments de deux de ces nouveaux ensembles est un problème encore plus difficile que ceux dont nous avons parlé tout à l'heure On s'en rendra compte aisément en cherchant à établir une telle correspondance entre les fonctions continues d'une variable réelle, d'une part, et les fonctions analytiques d'une variable complexe, d'autre part. A plus forte raison aura-t-on des difficultés inextricables si l'on veut introduire la considération de fonctions infiniment voisines.

Ces faits montrent combien la notion de puissance, malgré son utilité incontestable et son intérêt, est peu riche en conséquences, parce que c'est, au fond, une notion très superficielle. La puissance d'un ensemble concret est une de ses propriétés (1), mais c'est loin d'être la plus importante. Aussi, sans méconnaître l'intérêt de ce résultat, par exemple : *l'ensemble des fonctions analytiques de deux variables a même puissance*

---

(1) La théorie des *types ordinaux* paraît être autrement riche que la théorie des puissances; mais elle paraît aussi autrement difficile à édifier entièrement. Aussi ai-je renoncé à l'idée que j'avais eue primitivement d'y consacrer une Note; car j'ai craint d'être à la fois trop long et trop incomplet Pour donner quelque idée de l'objet de cette théorie, je citerai le théorème suivant, dû à M. G. Cantor : *Soient deux segments de droite, de longueur égale à l'unité et sur chacun d'eux un ensemble dénombrable partout dense; on peut toujours numéroter les éléments de chacun de ces ensembles de manière que la lettre a désignant les éléments du premier et la lettre b les éléments du second, la disposition relative de a_n et de a_m soit la même que la disposition relative de b_n et de b_m.*

Il est nécessaire, évidemment, si l'extrémité droite, par exemple, de l'un des segments appartient à l'ensemble correspondant, qu'il en soit de même de l'extrémité droite de l'autre segment. J'ajouterai qu'on peut, en modifiant légèrement la démonstration de M. Cantor, compléter son théorème en ajoutant que ε *étant un nombre positif quelconque donné à l'avance, on peut s'arranger de manière que, quels que soient m et n, les segments $a_m a_n$ et $b_m b_n$ vérifient les inégalités*

$$1 - \varepsilon < \frac{a_m a_n}{b_m b_n} < 1 + \varepsilon$$

mais, bien entendu, si ε devient de plus en plus petit, on pourra être obligé de modifier le numérotage à mesure que ε variera.

que l'ensemble des fonctions analytiques d'une variable, on doit avouer qu'il est bien moins fécond en applications qu'on ne pourrait le croire d'abord. Il peut cependant, joint aux remarques qui précèdent, nous aider à comprendre assez nettement le rôle des fonctions arbitraires en Analyse.

III. — La notion de fonction arbitraire.

Dans bien des problèmes qui se traduisent par des équations aux dérivées partielles, et, en particulier, dans bien des questions de Géométrie et de Physique mathématique, on est amené à dire que la solution générale dépend d'un certain nombre de fonctions arbitraires à un nombre déterminé d'arguments. Par exemple, la solution générale du problème de la déformation d'une surface renferme *deux* fonctions arbitraires d'*une* variable; la famille de Lamé la plus générale dépend de *trois* fonctions arbitraires de *deux* variables. Quel sens doit-on attribuer à cette « terminologie qui a quelque chose d'un peu indéterminé (1) » ? Si nous ne connaissions pas les résultats établis dans le paragraphe précédent, nous aurions pu être tentés de croire qu'elle exprime ceci : *L'ensemble des surfaces applicables sur une surface donnée a même puissance que l'ensemble des systèmes de deux fonctions arbitraires d'une variable.* Ces fonctions arbitraires doivent d'ailleurs être supposées *continues* (2) pour que la surface correspondante soit continue; et *analytiques*, si l'on veut borner le problème à la recherche des surfaces analytiques.

Cette proposition est incontestable, mais elle resterait exacte si nous remplacions dans son énoncé les mots *une variable* par les mots *deux variables*, ou *trois variables;* car l'ensemble des fonctions continues (ou analytiques) de deux ou trois variables a même puissance que l'ensemble des fonctions continues (ou analytiques) d'une variable. La notion de puissance ne suffit donc pas pour justifier la manière de parler qui nous occupe; nous allons en rechercher l'origine pour tâcher de mieux en comprendre le sens. Cette origine est évidemment dans les théorèmes d'existence des intégrales des équations aux dérivées partielles. Bornons-nous, pour plus de netteté, au cas des fonctions analytiques, dans lequel ces théorèmes d'existence ont été mieux étudiés; des remarques analogues s'appliqueraient, *mutatis mutandis*, si l'on considérait des fonctions de variables réelles satisfaisant, ainsi que leurs dérivées, à des conditions de continuité convenables.

---

(1) Darboux, *Leçons sur les systèmes orthogonaux et les coordonnées curvilignes*, t. I, p. 3.

(2) On doit aussi supposer vérifiées certaines conditions de continuité relativement à leurs dérivées. Mais nous n'avons pas à entrer ici dans ces détails.

Étant donnée une équation aux dérivées partielles, par exemple la suivante

$$\frac{\partial^2 z}{\partial x^2} = F\left(x, y, z, \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}, \frac{\partial^2 z}{\partial x\, \partial y}, \frac{\partial^2 z}{\partial y^2}\right), \tag{1}$$

comment a-t-on été amené à dire que l'intégrale générale de cette équation dépend de *deux* fonctions arbitraires d'*une* variable? Supposons, pour plus de netteté, que la fonction F soit une fonction analytique *entière* des variables dont elle dépend et donnons-nous pour $x = 0$ les développements en série de z et de $\frac{\partial z}{\partial x}$ suivant les puissances de y,

$$\left\{\begin{aligned} z &= A_0 + A_1 y + A_2 y^2 + \ldots, \\ \frac{\partial z}{\partial x} &= B_0 + B_1 y + B_2 y^2 + \ldots. \end{aligned}\right. \tag{2}$$

L'équation aux dérivées partielles (1) permettra alors de calculer toutes les dérivées partielles de z pour $x = 0$, $y = 0$ et, si l'on pose

$$\frac{1}{\alpha!\,\beta!}\left(\frac{\partial^{\alpha+\beta} z}{\partial x^\alpha\, \partial y^\beta}\right)_{x=y=0} = C_{\alpha,\beta},$$

on pourra écrire

$$z = \Sigma\Sigma\, C_{\alpha,\beta}\, x^\alpha y^\beta. \tag{3}$$

Les coefficients $C_{\alpha,\beta}$ sont des fonctions des A et des B, et d'ailleurs quels que soient les A et les B, le développement (3) satisfait formellement à l'équation (1); mais, d'après le théorème de Cauchy, il est convergent dans le cas et dans le cas seulement où les développements (2) le sont.

Il est clair que chaque système de séries (2) détermine une seule série (3) et réciproquement chaque série (3) détermine d'une manière unique [1] le système des séries (2). Donc, l'ensemble des séries (3) a même puissance que l'ensemble des séries (2), c'est-à-dire la puissance du continu; mais il est clair qu'on peut ranger les coefficients A et B de manière qu'ils soient les coefficients d'une série unique, à une ou à plusieurs variables. Si donc ils sont regardés comme des nombres tout à fait arbitraires, il n'y a aucune raison pour rattacher l'ensemble des séries (3) à l'ensemble des séries (2), plutôt qu'à l'ensemble des séries

$$\Sigma\Sigma\, M_{r,s}\, x^r y^s,$$

par exemple. Mais le théorème de Cauchy nous apprend, et c'est là le point essentiel, que, parmi les développements (3), *on obtient tous ceux qui sont convergents en choisissant les* A *et les* B *de manière que les*

---

[1] Il suffit, en effet, de faire $x = 0$ dans la série (3) et dans sa dérivée par rapport à x pour avoir les séries (2).

développements (2) *soient eux-mêmes convergents.* Nous arrivons donc au résultat suivant : *Les conditions de convergence du développement formel de l'intégrale générale* (développement formel qui renferme, quelle que soit la nature de l'équation aux dérivées partielles considérée, une infinité dénombrable de constantes arbitraires) *sont identiques aux conditions de convergence des développements en série d'un nombre déterminé de fonctions analytiques à un nombre déterminé d'arguments.*

Tel est le sens précis qu'on doit attribuer à l'expression : *dépendre de p fonctions arbitraires de n variables.* Il y aurait lieu de reprendre, à ce point de vue, l'étude des diverses questions dans lesquelles cette expression s'introduit, de manière à se rendre compte de la légitimité de son emploi.

Dans certains cas, dont plusieurs sont bien connus, il n'y a aucun inconvénient à substituer *une seule* fonction arbitraire d une variable à *plusieurs* fonctions arbitraires d'une variable (1) : mais la question n'a pas été étudiée en général ; nous nous contentons de la signaler.

---

(1) Parmi les procédés qu'on peut employer, en voici un qui est très général, dans le cas des fonctions de variables réelles : il suffit de partager le domaine dans lequel la fonction donnée est définie, en plusieurs autres (qui peuvent être en infinité dénombrable) ; on a ainsi autant de fonctions que de domaines et d'ailleurs, par une transformation simple, on peut étendre le domaine dans lequel chacune d'elles est définie.

Pour les fonctions de variable complexe développées en série de Taylor, on peut écrire $f(x) = \varphi(x^2) + x\psi(x^2)$ et faire correspondre à $f(x)$ les fonctions $\varphi(y)$ et $\psi(y)$. Ces divers procédés peuvent d'ailleurs être variés à l'infini.

NOTE IV.

LES POLÉMIQUES SUR LE TRANSFINI ET SUR LA DÉMONSTRATION DE M. ZERMELO.

I. — A propos de l' « infini nouveau » (1).

Lorsque M. G. Cantor, il y a une vingtaine d'années, publia ses idées sur la numération des nombres plus grands que l'infini, le public mathématique ne les accueillit pas sans quelque défiance. Néanmoins, bien des analystes furent frappés par la beauté de ces spéculations et ne se laissèrent pas arrêter par la forme, volontairement un peu paradoxale par endroits de l'exposition de M. Cantor. Aussi les applications de ces théories ne tardèrent pas à se multiplier; en même temps, l'attention de tous ceux qui s'intéressent aux principes de la science mathématique fut éveillée et les idées nouvelles sont maintenant, si l'on peut ainsi s'exprimer, devenues classiques, à la fois pour le public mathématique et pour le public philosophique (2).

Cependant, malgré les nombreux travaux auxquels elles ont donné lieu, il me semble qu'on n'a pas, du moins à ma connaissance, mis nettement en lumière ce qu'il y a de vraiment *nouveau* dans l'infini de M. Cantor, ou *transfini*. Je voudrais montrer qu'il se pose, pour les mathématiciens, la question de savoir s'il est légitime d'introduire *un nouveau principe d'induction* ou, si l'on préfère, *un nouveau mode de raisonnement*. Je voudrais, en même temps, signaler cette circonstance aux philosophes, dans l'espoir que l'*observation*, appliquée à la formation de ce principe, permette d'éclairer d'un jour nouveau les discussions *a priori* sur l'origine de principes analogues.

Je ne supposerai pas connues les théories de M. Cantor; je me dispen-

(1) *Revue philosophique*, 1899.

(2) Je n'ai point à faire ici une Bibliographie complète; comme indices de cette tendance des idées de M. Cantor à devenir classiques, citons simplement la Thèse de doctorat ès lettres de M. Couturat, où elles se trouvent clairement exposées, la Thèse de doctorat ès sciences de M. Baire, où elles sont utilisées presque à chaque page. Enfin, tout récemment, M. Evellin les a résumées et critiquées dans un remarquable article de la *Revue philosophique* auquel j'ai emprunté l'expression d'*infini nouveau*, mais sans y attacher le sens un peu ironique qu'elle a pour son auteur.

serai d'ailleurs de les exposer, n'ayant besoin, pour mon but, que d'un théorème de Paul du Bois-Reymond, dont j'indiquerai tout d'abord l'énoncé. Les personnes qui connaissent les idées de M. Cantor reconnaîtront aisément comment l'exposition pourrait être reprise en se plaçant au point de vue du savant professeur de Halle; mais la voie que j'ai suivie me paraît avoir l'avantage de donner aux raisonnements une forme plus concrète et, par suite, plus claire pour la majorité des esprits. J'espère d'ailleurs que ce mode d'exposition, dans lequel le nom de M. Cantor tiendra peu de place, n'empêchera personne de comprendre combien est grande l'importance de ses travaux et quelle influence leur lecture a eue sur le développement des idées que je vais tenter d'esquisser ici (¹).

I.

Un théorème de Paul du Bois-Reymond devant jouer un rôle tout à fait essentiel dans ce qui suit, je vais tout d'abord essayer d'expliquer aussi clairement que possible en quoi il consiste. Les personnes familières avec les Mathématiques pourront se contenter d'en lire l'énoncé qui termine ce paragraphe (²).

J'appellerai ici *nombres* exclusivement les nombres entiers positifs; la lettre x désignera l'un quelconque d'entre eux. Je désignerai par $f(x)$ un nombre qu'on sait calculer lorsque l'on connaît le nombre x; et $f(x)$ sera dit une fonction de x. Par exemple, supposons que l'on convienne de désigner par $f(x)$ le triple du nombre x; on aura ainsi défini une fonction particulière de x; si $x = 3$, $f(x) = 9$; si $x = 17$, $f(x) = 51$. On écrira d'ailleurs, comme on sait, $f(x) = 3x$. La fonction $f(x)$ dont nous venons de parler a la propriété suivante : si l'on donne à x des valeurs de plus en plus grandes, la fonction $f(x)$ prend elle-même des valeurs de plus en plus grandes; nous exprimerons ce fait en disant que la fonction $f(x)$ est *croissante*. Si l'on inscrit sur une ligne horizontale les valeurs des nombres entiers, dans leur ordre naturel, et, au-dessous de chacun d'eux, la valeur correspondante de la fonction $f(x)$, on pourra affirmer que, sur la seconde ligne comme sur la première, un nombre quelconque est supérieur à tous ceux qui sont à sa gauche, et inférieur à tous ceux qui sont à sa droite :

$$\begin{array}{llllllll} x & : 1, & 2, & 3, & 4, & 5, & \ldots, & 17, & \ldots; \\ f(x) & : 3, & 6, & 9, & 12, & 15, & \ldots, & 51, & \ldots. \end{array}$$

(¹) Je tiens d'autant plus à faire cette observation que, faute d'avoir insisté sur ce point dans un petit livre récent (c'est à la première édition de ce livre-ci qu'il est fait allusion), le choix d'un mode d'exposition autre que celui de M. Cantor a amené certains lecteurs à croire que je n'attachais pas une haute valeur à ses travaux. Le mérite de l'inventeur subsiste toujours même si, pour une raison ou une autre, la méthode qu'il a suivie pour arriver au but est abandonnée.

(²) *Voir* plus haut Note II.

En même temps que la fonction $f(x)$ que nous venons de considérer, et qu'on désigne par l'écriture $3x$, considérons la fonction $4x$; c'est-à-dire le quadruple de x. Il est clair que, quel que soit le nombre x, le nombre $4x$ est plus grand que le nombre $3x$; par exemple, si $x = 10$, $3x = 30$ et $4x = 40$. Nous exprimerons ce fait en disant que *la fonction croissante $4x$ est plus grande que la fonction croissante* $3x$. C'est une convention de langage. Considérons maintenant le carré de x, ou produit de x par lui-même; c'est aussi une fonction de x, qu'on désigne par x^2; nous allons la comparer avec la fonction $4x$. Pour cela nous écrirons sur une première ligne horizontale les valeurs successives du nombre x, puis, au-dessous, les valeurs correspondantes de $4x$ et, encore au-dessous, les valeurs correspondantes de x^2. Nous obtenons ainsi le Tableau suivant :

x :	1,	2,	3,	4,	5,	6,	7,	8,	...,	100,	...;
$4x$:	4,	8,	12,	16,	20,	24,	28,	32,	...,	400,	...;
x^2 :	1,	4,	9,	16,	25,	36,	49,	64,	...,	10000,	....

On voit que les nombres de la troisième ligne (ou valeurs de x^2) sont d'abord plus petits que ceux de la seconde (ou valeurs de $4x$), mais qu'ils deviennent ensuite plus grands, et l'on se rend compte sans peine que plus x devient grand, plus x^2 dépasse $4x$. Nous conviendrons encore de dire que *la fonction x^2 est plus grande que la fonction* $4x$, bien que le nombre x^2 ne soit pas toujours plus grand que le nombre $4x$; il nous suffit que cela ait lieu *à partir d'une valeur de x* (pour préciser ici, lorsque x dépasse 4). Cette définition étant admise, considérons la suite des fonctions croissantes qu'on obtient en multipliant x par les nombres entiers successifs, pris dans leur ordre naturel, c'est-à-dire la suite des fonctions :

$$(\alpha) \qquad x, \quad 2x, \quad 3x, \quad 4x, \quad 5x, \quad \ldots, \quad 100x, \quad \ldots.$$

Cette suite (α) doit être regardée comme indéfinie, de même que la suite naturelle des nombres entiers; elle a, de plus, une des propriétés capitales de cette dernière suite; chacune des fonctions (α) est plus grande que celles qui sont écrites à sa gauche. Mais voici une propriété de la suite (α) qui n'a pas d'analogue pour la suite naturelle des nombres entiers : *On peut trouver une fonction croissante plus grande que toutes les fonctions de la suite* (α). On constate, en effet, aisément, que la fonction x^2 a cette propriété. Il importe ici d'éviter toute confusion. Lorsqu'on donne à x une valeur déterminée, par exemple, $x = 50$, x^2 a une valeur déterminée (ici 2500); d'autre part, les fonctions (α) ont une suite de valeurs de plus en plus grandes,

$$50, \quad 100, \quad 150, \quad 200, \quad 250, \quad \ldots, \quad 5000, \quad \ldots,$$

parmi lesquelles il y en a une infinité dépassant la valeur de x^2. Mais donnons-nous à l'avance une fonction de la suite (α) c'est-à-dire, choisis-

sons dans cette suite une fonction déterminée, par exemple $1000x$. *Quelle que soit la manière dont nous l'ayons choisie, nous pourrons affirmer que la fonction x^2 est plus grande*, c'est-à-dire la dépasse à partir d'une certaine valeur de x (ici, à partir de $x = 1000$). Cette valeur de x, il est vrai, varie lorsque la fonction choisie varie elle-même; il nous suffit qu'elle existe toujours pour que notre énoncé : *la fonction x^2 est plus grande que toutes les fonctions* (α), ait un sens bien précis.

Supposons maintenant que nous ayons une suite indéfinie de fonctions de x analogue à la suite (α), c'est-à-dire telle que chaque fonction de la suite soit plus grande que les précédentes.

Soit :

$$(\beta) \qquad f_1(x), \quad f_2(x), \quad f_3(x), \quad f_4(x), \quad \ldots, \quad f_{100}(x), \quad \ldots$$

cette suite; le théorème de Paul du Bois-Reymond est le suivant : *Il existe une fonction plus grande que toutes les fonctions de la suite* (β). De plus, la démonstration, que nous omettons, indique le moyen de trouver cette fonction. Par exemple, si les fonctions (β) sont les puissances successives de x, c'est-à-dire si l'on a

$$f_1(x) = x, \quad f_2(x) = x \times x = x^2, \quad f_3(x) = x \times x \times x = x^3,$$
$$f_4(x) = x \times x \times x \times x = x^4, \quad f_5(x) = x \times x \times x \times x \times x = x^5, \quad \ldots,$$

on verra facilement que la fonction 10^x, c'est-à-dire le produit de x facteurs égaux à 10 satisfait à la condition requise, c'est-à-dire est plus grande que chacune des fonctions

$$x, \quad x^2, \quad x^3, \quad x^4, \quad x^5, \quad \ldots, \quad x^{100}, \quad \ldots.$$

Pour se convaincre, par exemple, que la fonction 10^x est plus grande que x^{100}, il suffit de remarquer que, pour $x = 1000$, 10^x est égal à l'unité suivie de 1000 zéros, tandis que x^{100} est égal seulement à l'unité suivie de 300 zéros.

Mais ce résultat particulier n'est pas ce qui nous importe; ce qui fait le plus grand intérêt du théorème de Paul du Bois-Reymond, c'est qu'il s'applique, *quelle que soit la suite* (β). On peut l'énoncer en disant que : *Étant donnée une suite indéfinie quelconque de fonctions croissantes*[1], *il existe une fonction croissante plus grande que chacune d'elles.* C'est sur cet énoncé que reposent toutes les considérations qui suivent; j'espère l'avoir rendu clair.

II.

Il n'est cependant pas inutile de préciser un peu le sens d'un mot qui figure dans notre énoncé, celui de la suite *indéfinie*. Nous avons, dans le

---

[1] Il n'est même pas nécessaire de supposer les fonctions rangées de telle manière que chacune soit plus grande que les précédentes.

paragraphe précédent, considéré des suites telle que (α), (β), se terminant par des points de suspension ; la signification d'une telle écriture était d'ailleurs sans doute claire pour tous nos lecteurs; les suites ainsi écrites doivent être considérées comme prolongées indéfiniment, *à la manière de la suite naturelle des nombres entiers*. La conception que nous en avons est donc précisément aussi claire, ni plus, ni moins, que notre conception de cette suite natutelle. Ce n'est point ici le lieu de discuter l'origine ni les fondements de cette conception; indiquons seulement, ce point ayant pour nous grande importance, que beaucoup de questions relatives à la suite naturelle des nombres entiers se ramènent à la suivante : la proposition fondamentale *après tout nombre entier il y en a un autre*, suffit-elle à nous donner la conception de cette suite ? ou, en d'autres termes, cette conception renferme-t-elle, ou non, quelque chose de plus que la proposition fondamentale ?

Mais revenons à la suite (β); quand on veut préciser la propriété qu'elle possède de *ressembler* à la suite naturelle des nombres entiers, on dit souvent qu'elle est *dénombrable;* cela signifie qu'on peut attribuer à chacune des fonctions qui la composent un numéro d'ordre, et *dénombrer* ainsi l'ensemble de ces fonctions au moyen des nombres entiers sucessifs :

$$1,\quad 2,\quad 3,\quad 4,\quad 5,\quad \ldots,\quad 100,\quad \ldots.$$

L'étude des suites dénombrables est fort importante en Mathématiques; parmi les propositions qu'on rencontre dans cette étude, l'une des premieres, et non la moins importante est la suivante : *Lorsqu'on a plusieurs suites dénombrables, on peut ranger leurs éléments de manière à n'en former qu'une seule.* Par exemple, si l'on a les deux suites

$$\begin{array}{llllllll} a_1, & a_2, & a_3, & a_4, & \ldots, & a_{100}, & \ldots, \\ b_1, & b_2, & b_3, & b_4, & \ldots, & b_{100}, & \ldots, \end{array}$$

il suffira d'écrire comme il suit les éléments dans une suite unique :

$$\begin{array}{llllllllll} a_1, & b_1, & a_2, & b_2, & a_3, & b_3, & \ldots, & a_{100}, & b_{100}, & \ldots, \\ 1, & 2, & 3, & 4, & 5, & 6, & \ldots, & 199, & 200, & \ldots; \end{array}$$

au-dessous de chaque élément, nous avons écrit le numéro d'ordre qu'il prend dans la nouvelle suite.

III.

Ces préliminaires établis, nous pouvons aborder le problème suivant : *Chercher à former une suite de fonctions croissantes de plus en plus grandes.* Remarquons tout de suite l'analogie de ce problème avec celui qui se pose au début de la numération : *Former des nombres entiers de plus en plus grands.* Ce dernier problème se résout à l'aide du principe que nous rappelions tout à l'heure : *Après tout entier, il y en a un*

autre; c'est le théorème de Paul du Bois-Reymond qui, *dans la numération des fonctions croissantes*, remplacera pour nous ce principe.

Voici donc comment nous procéderons; nous partirons d'une fonction croissante déterminée, par exemple x^2; il n'est pas difficile, en partant de cette fonction, de former une suite indéfinie de fonctions croissantes de plus en plus grandes [1]. Cette suite étant obtenue, le théorème de Paul du Bois-Reymond nous fournit une fonction plus grande que chacune des précédentes; nous pouvons partir de cette nouvelle fonction pour former une suite indéfinie de fonctions de plus en plus grandes, appliquer à cette nouvelle suite le théorème, partir de la fonction ainsi obtenue pour former une nouvelle suite, et continuer ainsi indéfiniment.

D'ailleurs, continuer indéfiniment signifie que nous aurons à appliquer notre théorème une première fois, puis une seconde, puis une troisième, etc.; nous sommes conduits chaque fois à une fonction plus grande que toutes les précédentes. Or, les fonctions ainsi obtenues forment une suite dénombrable; nous pouvons donc appliquer encore à cette suite le théorème, et obtenir ainsi une fonction encore plus grande, etc. On voit que l'application de ces procédés *n'a pas de limite*, mais il importe de bien comprendre ce qu'il faut entendre par là, car c'est le point capital dans cet article. Lorsque nous appliquons un procédé, quel qu'il soit, *indéfiniment*, nous entendons par là que nous l'appliquons une fois, deux fois, trois fois, etc., bref, aussi longtemps qu'il y a des nombres entiers. Les résultats obtenus se présentent donc sous la forme d'une suite analogue à la suite des nombres entiers; nous n'en atteignons pas l'extrémité, mais nous avons, ou croyons avoir, une conception nette d'une telle suite.

Dans la numération des fonctions croissantes, nous ne pourrons jamais nous arrêter à une telle suite; car, d'après le théorème fondamental, lorsque nous aurons une telle suite dénombrable de fonctions, nous pourrons trouver une fonction plus grande [2]. De même que, étant donné un nombre limité de nombres entiers, il existe un nombre entier plus grand que chacun d'eux; de même, étant donnée une suite dénombrable quelconque de fonctions croissantes, il existe une fonction croissante plus grande que chacune d'elles. On peut dire aussi que la numération des entiers n'est pas *limitée*, et que la numération des fonctions croissantes

---

[1] L'un des procédés les plus commodes pour arriver à ce but est l'*itération*, c'est-à-dire la répétition indéfinie de la même opération. Notre première fonction étant x^2, c'est-à-dire le carré de x, nous pourrons prendre pour la seconde le carré du carré de x, pour la troisième le carré du carré du carré de x, pour la quatrième le carré du carré du carré du carré de x, et ainsi de suite. Il est clair qu'on obtient ainsi des fonctions de plus en plus grandes et qu'on peut continuer *indéfiniment*, c'est-à-dire aussi longtemps qu'il y a des nombres entiers.

[2] Pour démontrer rigoureusement ce dernier résultat, il faut se servir de cette proposition, d'ailleurs aisée à établir, qu'un ensemble dénombrable d'ensembles dénombrables est lui-même dénombrable. Mais ce sont là des questions techniques ici sans importance.

n'est pas *dénombrable*, c'est-à-dire ne peut être réalisée au moyen de suites indéfinies analogues à la suite naturelle des nombres.

Voici maintenant la question qui se trouve posée aux mathématiciens; nous tâcherons d'indiquer ensuite quel intérêt elle a pour les philosophes.

L'un des procédés de raisonnement les plus employés en Mathématiques est l'induction, par passage de n à $n+1$. On vérifie qu'une proposition est exacte si les nombres entiers qui figurent dans son énoncé sont assez petits; on démontre ensuite qu'elle subsiste si l'un quelconque de ces nombres augmente d'une unité et l'on en conclut qu'elle est vraie pour des entiers *quelconques*, car par des augmentations successives d'une unité, un nombre entier peut devenir aussi grand qu'on veut [1]; c'est d'ailleurs dans l'emploi du mot *quelconques* que réside l'induction: c'est ce mot qui transforme la proposition particulière dont on est parti en une proposition générale.

Cela étant rappelé, la question dont nous voulions parler est la suivante : *Y a-t-il lieu d'admettre un nouveau mode d'induction s'étendant, non plus seulement à la suite dénombrable des nombres entiers, mais à la suite non dénombrable des fonctions à croissance de plus en plus grande?*

On peut donner d'ailleurs à cette question des formes très diverses, mais au fond équivalentes. On peut, par exemple, demander simplement si nous pouvons avoir la conception de la suite non dénombrable des fonctions croissantes de plus en plus grandes. La difficulté qu'il y a à acquérir cette conception tient à ce que tout procédé *déterminé* de formation de cette suite, par cela seul qu'il s'exprime au moyen d'un nombre fini de mots, ne conduit qu'à une suite dénombrable; mais, d'autre part, le théorème de Paul du Bois-Reymond nous apprend à dépasser cette suite dénombrable une fois formée. Il y a là une antinomie que l'analyse paraît impuissante à résoudre; il ne paraît cependant pas douteux que le développement ultérieur des Mathématiques ne conduise nécessairement à faire un choix, c'est-à-dire à admettre ou à proscrire le *transfini* (ou infini non dénombrable).

Si on le proscrit, pour la raison que nous ne pouvons en avoir une conception claire, ce sera un argument très fort pour ceux qui prétendent que notre conception de la suite indéfinie des nombres entiers renferme quelque chose en plus de la simple remarque : Après tout entier, il y en a un autre. Car si cette remarque suffisait à donner la notion de l'infini dénombrable, le théorème de Paul du Bois-Reymond devrait suffire à nous donner la notion du transfini.

Si, au contraire, on admet la notion du transfini, on se trouve avoir créé véritablement un *infini nouveau* et l'étude des procédés par lesquels cette *création* d'une notion nouvelle aura été obtenue et légitimée paraît pouvoir présenter de l'intérêt pour le philosophe.

---

[1] *Cf.* H. POINCARÉ, *De la nature du raisonnement mathématique* (*Revue de Métaphysique et de Morale*, 1892).

Je me contenterai de ces indications, ne voulant pas aborder le fond de la question; d'ailleurs, à mon sens, telle ou telle opinion personnelle ne peut avoir, en cette matière, une très grande importance. Ce qui est important, c'est la solution que les progrès de la Science amèneront les mathématiciens à adopter unanimement : car se sera *un fait*.

II. — L'antinomie du transfini (1).

La lecture de l'article publié récemment ici même par MM. Evellin et Z. (2) m'a montré que, dans les quelques pages que j'ai écrites sur l'*infini nouveau* (3) ou *transfini*, le passage le plus important n'est pas suffisamment clair. Il me paraît donc utile de chercher à élucider définitivement la manière dont se présente, en *Mathématiques*, la notion du transfini; j'ajouterai ensuite quelques observations sur les conséquences qui me paraissent en découler relativement à l'origine psychologique de la notion de l'indéfini. Cette seconde Partie, qui n'avait pas d'analogue dans mon premier article, écrit à un point de vue exclusivement mathématique, pourra peut-être utilement donner lieu à discussion.

I.

Rappelons d'abord l'énoncé d'un théorème de Paul du Bois-Reymond, qui joue un rôle essentiel dans tout ceci et que nous appellerons, pour abréger, *le théorème fondamental : Etant donné un ensemble dénombrable quelconque de fonctions croissantes, il existe une fonction croissante plus grande que chacune d'elles*. D'ailleurs, parmi les différents termes qui figurent dans cet énoncé, les seuls dont il soit essentiel de connaître le sens précis sont ceux d'*ensemble dénombrable;* il est inutile de rappeler les définitions des autres; le lecteur pourra les trouver dans les articles déjà cités. Mais, pour éviter toute confusion, rappelons qu'un ensemble *dénombrable* est un ensemble tel qu'on puisse faire correspondre ses éléments aux nombres entiers positifs, de telle manière qu'à chaque élément corresponde *un* entier positif déterminé et à chaque entier positif *un* élément déterminé. D'ailleurs, on dira qu'un tel ensemble est *donné* lorsque, si l'on prend au hasard un entier positif *quelconque*, par exemple 22315, il est possible, par des opérations en nombre limité, de déterminer d'une manière précise l'élément de l'ensemble qui correspond au nombre 22315.

Cherchons maintenant à former une suite de fonctions croissantes de plus

(1) *Revue philosophique*, 1900.
(2) *Ibid.*, février 1900.
(3) *Ibid.*, octobre 1899 (reproduit dans le premier paragraphe de cette Note).

en plus grandes. En partant d'une fonction déterminée, on peut aisément, par l'itération, comme je l'ai expliqué dans mon premier article, obtenir une première suite dénombrable de fonctions de plus en plus grandes. En appliquant à cette première suite le théorème fondamental, on obtient une fonction plus grande que les précédentes et, en la prenant comme « terme de tête », on en déduit une seconde suite dénombrable, à laquelle on applique de nouveau le théorème fondamental, ce qui donne une nouvelle fonction plus grande que les précédentes et qui peut servir de nouveau terme de tête, etc.

Si d'ailleurs on entend par « *etc.* » qu'on répète la même opération une infinité dénombrable de fois, on obtiendra une infinité dénombrable de *termes de tête* et l'ensemble total des fonctions obtenues sera dénombrable, en vertu d'un théorème connu que j'avais seulement énoncé dans mon article, mais dont MM. Evellin et Z. ont donné une démonstration très claire [1].

Nous avons donc obtenu un ensemble dénombrable de fonctions croissantes; cet ensemble est d'ailleurs *donné*, au sens précisé plus haut. Dès lors, *et c'est là le point capital*, nous pouvons appliquer à cet ensemble le théorème fondamental, c'est-à-dire trouver une fonction plus grande que toutes les précédentes. Il n'est donc pas possible de se borner à considérer ces dernières.

Il est inutile de fatiguer le lecteur en répétant toujours les mêmes mots : l'application des procédés précédents est indéfinie. D'ailleurs, tels procédés déterminés qu'on voudra, appliqués à une infinité dénombrable de fois, ne conduiront jamais qu'à un ensemble dénombrable. Seulement, *une fois cet ensemble obtenu*, on pourra le dépasser par une application nouvelle du théorème fondamental.

Il y a là une antinomie tout à fait analogue à celle qui se présente lorsqu'on cherche à former les nombres entiers successifs; en appliquant un nombre *limité* de fois un mode quelconque de formation, on n'aura jamais qu'un nombre *limité* d'entiers : on ne peut pas créer l'illimité avec le limité, l'infini avec le fini. Mais cependant, on ne peut pas dire que la suite des nombres entiers soit limitée, puisque après tout entier il y en a un autre.

De même, il n'est pas possible de créer le non-dénombrable avec le dénombrable (ou, si l'on veut, le transfini avec l'indéfini). Mais cependant on ne peut pas dire que la suite des fonctions croissantes soit dénombrable, puisque le théorème fondamental permet de dépasser toute suite dénombrable à laquelle on pourrait être tenté de se borner.

Les mathématiciens ont donné depuis longtemps, à l'antinomie relative à la suite des nombres entiers, une solution *de fait*, complètement indépendante des questions métaphysiques qui peuvent y être liées. Ils se servent de la suite indéfinie des nombres entiers, comme si cette suite

(1) *Revue philosophique*, 1900, p. 137.

existait réellement dans son ensemble [1] et, pour que cette manière de procéder soit légitime, il suffit qu'elle leur rende des services, ce que nul ne peut nier.

De même, le développement ultérieur des Mathématiques conduira *peut-être* à introduire dans les raisonnements une suite transfinie de fonctions croissantes de plus en plus grandes. A partir de ce moment, le transfini aura conquis droit de cité en Mathématiques, au même titre que l'indéfini.

Sans vouloir trancher la question soulevée par ce *peut-être* on peut faire les remarques suivantes : Si la proposition « après tout entier, il y en a un autre » suffit à légitimer l'introduction de l'indéfini, le théorème fondamental doit suffire à légitimer l'introduction du transfini. On a donc le choix entre deux alternatives : proscrire le transfini, mais admettre que la notion de l'indéfini, telle qu'elle est utilisée en Mathématiques, renferme quelque chose de plus que la proposition « après tout entier il y en a un autre » ; ou bien, niant ce dernier point, admettre un *infini nouveau*, le transfini essentiellement distinct de l'indéfini.

II.

J'abandonne maintenant le terrain du raisonnement mathématique pour celui de la controverse philosophique, qui est loin de m'être aussi familier, comme on s'en apercevra sans doute. Pour éviter toute confusion, je crois devoir commencer par déclarer que je ne m'occupe nullement de ce qu'on peut appeler la *notion métaphysique de l'infini,* c'est-à-dire de la question de savoir s'il peut correspondre quelque réalité à notre concept de l'infini. Je voudrais me contenter de dire quelques mots d'un problème purement psychologique, qui peut s'énoncer ainsi : la notion de l'indéfini, ou infini dénombrable, est, pour les mathématiciens, une notion parfaitement claire et ne donnant lieu, lorsqu'ils parlent entre eux, à aucune ambiguïté; il n'y a donc pas lieu de discuter la vérité de cette notion, en tant que notion; elle est vraie par cela seul qu'elle existe; mais on peut rechercher l'origine psychologique de son existence, c'est-à-dire chercher à l'expliquer par des notions plus simples ou tout au moins plus primitives.

Il n'est pas contestable que la remarque : *après tout entier il y en a un autre*, manifestement antérieure à l'acquisition de la notion de l'indéfini, n'ait joué un rôle essentiel dans la formation de cette notion. Toute la question est de savoir si elle suffit, *à elle seule*, à expliquer cette formation, ou s'il faut y ajouter quelque chose, et, dans ce dernier cas, quelle chose ?

---

[1] Par exemple, on dira que l'équation $\sin \pi x = 0$ admet comme racine *tous* les nombres entiers, et la considération *simultanée* de toutes ces racines donne l'expression de la fonction sous forme de produit infini.

On pourrait faire observer que, soit dans l'histoire de la Science, soit dans celle de chaque esprit, il s'écoule, en fait, un temps considérable entre le moment où la remarque peut être faite et celui où la notion de l'indéfini peut être comprise. Mais il est facile d'objecter à cela que, si quelque réflexion et quelque maturité d'esprit sont nécessaires pour acquérir la notion, après avoir fait la remarque, cela ne suffit pas à prouver que cette acquisition exige l'addition d'un élément étranger. N'insistons donc point sur cet argument.

Mais on peut tirer de notre étude du transfini un argument plus sérieux : il ne semble pas, en effet, que la connaissance du théorème fondamental suffise pour donner une notion nette du transfini. Du moins, tous ceux dont l'instruction mathématique est assez avancée pour qu'ils comprennent ce théorème aussi clairement que chacun comprend qu'après chaque entier il y en a un autre, n'ont pas en commun une notion du transfini, aussi claire que la notion qu'ils ont de l'indéfini, c'est-à-dire telle qu'ils puissent se comprendre entre eux lorsqu'ils en parlent, sans aucune chance d'erreur. Par exemple, les uns consentiront volontiers à raisonner sur des suites transfinies réalisées, de même qu on raisonne à chaque instant sur des suites indéfinies réalisées (indépendamment de toute hypothèse métaphysique, répétons-le), tandis que d'autres se refuseront à ces raisonnements.

La question est maintenant de savoir si ces derniers sont dans leur droit, ou bien s'ils se trouvent vis-à-vis du transfini dans une situation analogue à celle d'un enfant qui vient d'apprendre la numération, à qui on expliquerait ce qu'est un ensemble dénombrable et qui aurait quelque peine à le comprendre. Cette dernière hypothèse doit être adoptée par ceux qui prétendent que la notion de l'indéfini ne contient rien de plus que la remarque : après tout entier il y en a un autre; ils doivent admettre aussi que le théorème fondamental suffit à donner la notion du transfini. Sans prétendre trancher une question qui, à mon avis, ne sera résolue définitivement que par l'observation des faits, c'est-à-dire par l'étude du développement ultérieur des Mathématiques, je voudrais indiquer les raisons pour lesquelles ceux qui se refusent à considérer la notion du transfini comme *actuellement aussi claire* que celle de l'indéfini me paraissent être dans le vrai. Je ne nie pas d'ailleurs qu'elle ne puisse le devenir; si ce résultat est atteint un jour, l'étude des procédés par lesquels l'esprit sera arrivé à ce résultat me paraît devoir être très intéressante; il n'y a, en effet, peut-être pas d'exemple d'une autre circonstance où l'on ait pu ainsi observer le fonctionnement de l'esprit cherchant à acquérir une notion abstraite nouvelle, et à créer en même temps un mode de raisonnement nouveau.

Que faut-il donc ajouter à la remarque « après tout entier il y en a un autre » pour acquérir la notion de l'indéfini ? Ceci, je crois : l'opération par laquelle l'addition d'une unité permet avec chaque entier de former l'entier suivant peut être indéfiniment considérée comme étant, à un certain point de vue, *la même*. Il est clair, en effet, que ce n'est pas, *en soi*,

la même chose d'ajouter 1 à 1 pour obtenir 2, ou d'ajouter 1 à 2 pour obtenir 3. On ne peut même pas prétendre qu'on est *naturellement* conduit à regarder comme identiques l'opération par laquelle on passe de 1 à 2 et celle par laquelle on passe de 2 à 3; tout dépend du point de vue sous lequel on le considère ([1]). D'autre part, étant donnée une infinité (dénombrable) d'opérations telles que chacune conduise d'un nombre donné à un autre nombre donné, on prouve aisément qu'on peut trouver une formule mathématique assez générale pour les donner toutes, c'est-à-dire que, en un certain sens, on peut considérer toutes ces opérations comme identiques. Mais ce résultat n'est obtenu que par des raisonnements compliqués, tandis que l'esprit aperçoit immédiatement que l'opération par laquelle on passe d'un entier à l'entier suivant peut, à un certain point de vue, être indéfiniment considérée comme la même.

La question n'est d'ailleurs pas résolue complètement, puisque nous avons dû faire entrer le mot *indéfiniment* dans notre explication; elle est seulement déplacée. Il nous faut chercher à comprendre comment, avant d'avoir l'idée de l'indéfini, on peut concevoir comme identiques des opérations en nombre indéfini. Cela ne peut évidemment se produire que par la connaissance d'une propriété commune à ces opérations, propriété indépendante de leur nombre.

C'est donc dans les propriétés fondamentales de l'*addition* (commutativité, etc.) que nous trouvons l'origine de la notion d'*indéfini*. Nous ne nous étendrons pas sur les développements mathématiques qui trouveraient ici leur place; rappelons seulement les travaux de M. Drach, et notamment sa Thèse ([2]), dans laquelle il montre comment on peut construire les Mathématiques en partant seulement des *propriétés* des opérations fondamentales : addition, multiplication, différentiation.

Il ne reste qu'une difficulté à résoudre : Comment s'aperçoit-on que l'opération *addition*, définie par ses propriétés fondamentales, est la même que l'opération *addition* par laquelle on passe de chaque entier au suivant ? Nous nous contenterons de dire que c'est une vérité qui s'impose à l'esprit de chacun de nous; ce serait sortir de notre sujet que d'intervenir dans les discussions, de nature métaphysique, auxquelles la question de savoir pourquoi il en est ainsi a depuis longtemps donné lieu, sous une forme à peine différente.

Contentons-nous d'avoir constaté que, pour acquérir la notion de l'indéfini, telle qu'on en use en Mathématiques, il faut ajouter quelque chose à

([1]) Par exemple, si l'on a trois cordes, telles que la seconde effectue deux vibrations pendant que la première en effectue une, tandis que la troisième en effectue trois, tandis que la seconde en effectue deux, on sait que la seconde donne l'octave de la première, tandis que la troisième donne la quinte de la seconde. Il n'est donc pas *naturel* de dire qu'il y a le même intervalle entre les deux premières qu'entre les deux dernières.

([2]) *Essai sur une théorie générale de l'intégration et sur la classification des transcendantes* (*Annales de l'École Normale*, 1898).

la remarque d'après laquelle chaque entier est suivi d'un autre. On voit dès lors où gît la difficulté dans l'antinomie du transfini; il s'agit d'arriver à comprendre nettement *à quel point de vue* les applications successives qu'on fait du théorème fondamental peuvent être regardées comme étant *la même* opération [1]. De même, dans l'étude du continu, il n'y a aucune difficulté tant qu'on reste au point de vue purement géométrique; les points d'une droite, par exemple, sont tous identiques, car ils ont tous les mêmes propriétés générales; les difficultés ne se présentent qu'avec les définitions arithmétiques, car la connaissance de ces propriétés générales communes est alors moins aisée.

III. — Le continu bien ordonné d'après M. Zermelo.

On dit qu'un ensemble est ordonné, si l'on a établi entre ses éléments des relations telles que, a et b, étant deux éléments quelconques, on puisse dire que l'un d'eux a est *avant* b, ce qu'on écrira $a < b$ ou $b > a$; les relations $a < b$ et $b < c$ doivent entraîner $a < c$. On dit en outre que l'ensemble E est *bien ordonné* si, étant donné un sous-ensemble quelconque E′, il y a dans E′ un élément a situé avant tous les autres éléments de E′.

L'ensemble des nombres compris entre 0 et 1 est ordonné puisqu'on peut convenir de regarder a comme situé avant b si a est plus petit que b, au sens arithmétique; mais cet ensemble n'est pas bien ordonné, car, si l'on considère les nombres *supérieurs* à 0,5 comme formant un ensemble E′, il n'y a pas d'élément de E′ qui soit inférieur à tous les autres.

Le type des ensembles bien ordonnés est fourni par les nombres transfinis de M. Cantor; nous avons indiqué comment on peut définir les premiers nombres de la seconde classe et essayer de concevoir les autres nombres de la seconde classe; certains théoriciens des ensembles vont plus loin et désignent par la lettre hébraïque aleph ($\aleph$) affectée d'indices successifs, les puissances des classes successives. Aleph-zéro est alors la puissance des ensembles dénombrables (leur nombre cardinal), aleph-un, la puissance des nombres de la seconde classe de M. Cantor, etc.

La question de savoir si la puissance du continu est exprimable par un aleph (d'indice connu ou inconnu) a été posée par M. Cantor depuis longtemps; elle paraît équivalente au problème suivant : le continu peut-il être bien ordonné ?

---

(1) *Voyez* l'article bien connu de M. Poincaré : *De la nature du raisonnement mathématique* (*Revue de Métaphysique et de Morale*, 1893). La considération de l'indéfini y est ramenée à une *induction* d'une nature particulière qui constitue l'essence même du raisonnement mathématique. La question que nous avons posée se ramène alors à celle-ci : doit-on admettre une nouvelle induction, s'étendant, non plus seulement jusqu'à l'infini, mais jusqu'au transfini ?

D'un point de vue idéaliste, on peut concevoir que tout ensemble puisse être bien ordonné, puisqu'il suffit de choisir successivement dans cet ensemble des éléments auxquels on assignera comme rangs les nombres transfinis successifs; si l'on admet que la numération des nombres transfinis n'a pas de limite, on peut concevoir que l'on arrive ainsi, au moins idéalement, à épuiser tout ensemble. Les objections contre cette « démonstration » ont déjà été discutées et il est inutile d'y insister à nouveau. C'est pour y échapper que M. Zermelo a donné une autre démonstration du fait que tout ensemble peut être bien ordonné; la démonstration de M. Zermelo remplace les choix successifs dont il vient d'être question par des choix indépendants, pouvant ainsi être conçus comme simultanés; mais nous allons voir que ces choix sont « transfiniment » plus nombreux et que, en réalité, la démonstration de M. Zermelo n'est qu'une complication de la démonstration intuitive qui repose sur des choix successifs, complication dissimulée sous un appareil logique.

J'emprunterai les grandes lignes de la démonstration de M. Zermelo à l'exposition de M. Hausdorff [1], en essayant de la simplifier le plus possible et de la traduire en langage ordinaire.

Il s'agit de prouver qu'un ensemble quelconque M peut être bien ordonné. Soit A un sous-ensemble quelconque de M; nous admettons qu'on choisit dans chaque A (M compris) un élément a qui sera dit l'*élément distingué de* A. L'ensemble A′ formé des éléments de A, sauf a, est dit le *suivant de* A.

On dira qu'un ensemble E de sous-ensembles tels que A forme une *chaîne* si : 1° M fait partie de E; 2° le suivant A′ de tout ensemble A de E appartenant aussi à E; 3° le diviseur d'un ensemble quelconque d'ensembles de E appartient à E (le diviseur d'un ensemble fini ou infini d'ensembles A est l'ensemble des éléments communs à tous les A). Il est clair qu'il y a des *chaînes*, puisque l'ensemble de *tous* les sous-ensembles de M (M compris) en constitue une; il est évident que si l'on considère l'ensemble E de toutes les chaînes, l'ensemble diviseur de E est aussi une chaîne; c'est *la plus petite chaîne;* nous la désignerons par C. On va prouver que tous les éléments de C sont *normaux*, c'est-à-dire que A et B étant deux éléments quelconques de C, on a ou $A > B$, ou $A = B$, ou $A < B$. Cela résultera du fait que les éléments normaux de C forment une chaîne C′ qui est inférieure ou égale à C; comme C est la plus petite chaîne, C′ est identique à C. Appelons donc *élément normal de* C tout ensemble A comparable à tous les autres éléments de C; il y a des éléments normaux, car M fait partie de C (d'après la définition des chaînes) et est normal. D'autre part, si A est normal, son suivant A′ est normal; sinon l'ensemble des éléments de C supérieurs à A et l'ensemble des éléments de C inférieurs à A′ formeraient, en y joignant A et A′, une

[1] Hausdorff, *Grundzüge der Mengenlehre* (Leipzig, Veit, 1914).

chaîne C′ plus petite que C [1]. Enfin, le diviseur D d'ensembles normaux A est normal; car si B est inférieur ou égal à tous les A, il est inférieur ou égal à D, et si B est supérieur à l'un des A, il est supérieur à D. L'ensemble des éléments normaux de C forme donc une chaîne, forcément identique à C, c'est-à-dire que tous les éléments de C sont normaux. C est donc ordonné; M sera le premier élément, et A sera avant B si A est supérieur [2] à B; de plus, C est *bien ordonné*, car si nous considérons un sous-ensemble quelconque V de C, l'ensemble U des éléments de C supérieurs à tous les éléments de V admet un diviseur D qui fait partie de C, d'après la définition de la chaîne; tout élément de V est inférieur ou égal à D, et tout élément de U est supérieur ou égal à D; D est d'ailleurs le plus grand ensemble qui satisfasse à ces conditions; on en conclut immédiatement que D *ou* D′ est le premier élément de V.

L'ordination de C entraîne celle de M, car il y a entre les éléments de C et ceux de M une correspondance univoque et réciproque, à chaque ensemble A correspondant l'élément distingué a; en effet, si B diffère de A, on aura par exemple $B < A$ et, par suite, $B \leqq A'$; donc b ne peut pas être identique à a puisque A′ ne renferme pas a. D'autre part, à chaque élément a, on peut faire correspondre le diviseur A de tous les ensembles de C qui contiennent a; il est clair que a est l'élément distingué de A, sinon A′ contiendrait a et le diviseur des ensembles contenant a serait inférieur ou égal à A′, contrairement à l'hypothèse.

L'ensemble M est donc bien ordonné.

La méthode d'ordination intuitive que nous avons d'abord esquissée aurait conduit à choisir un élément m dans M, puis un élément m' dans M′, un élément m'' dans le suivant M″ de M′ et ainsi de suite transfiniment. Pour s'éviter ces choix successifs, M. Zermelo choisit un élément distingué, non seulement dans M′, mais dans *tous* les ensembles M_1 qui ne diffèrent de M que par un seul élément; de même, afin que le choix de m'' soit indépendant des choix de m et de m', il choisit un élément distingué m_2 dans tous les ensembles M_2 qui ne diffèrent de M que par deux éléments. Le nombre des choix à faire est d'une puissance supérieure (*voir* le § 2 de la Note I), mais ces choix peuvent être faits dans un ordre quelconque, ceux qui concernent les ensembles M_2 par exemple avant ceux qui concernent les ensembles M_1. Il s'agit ensuite de former toutes les chaînes et d'en déduire la plus petite chaîne C qui est leur plus petit diviseur commun. Le moyen qui me paraît le plus simple

---

[1] Ce point est à peu près évident; pour le démontrer en toute rigueur, il faut observer que si un ensemble B est supérieur à A, le suivant B′ de B ne peut pas être inférieur à A, car B′ ne diffère de B que par *un* élément, tandis qu'un ensemble supérieur à A et un ensemble inférieur à A diffèrent par au moins *deux* éléments; comme A est normal, B′ est supérieur ou égal à A.

[2] Il faut prendre garde que le nombre ordinal de A est inférieur à celui de B; le plus grand ensemble M est classé le *premier*, c'est-à-dire a le plus petit nombre ordinal.

pour y arriver, c'est d'observer que l'ensemble formé par M, M' et les sous-ensembles de M' constitue une chaîne; la plus petite chaîne ne contient donc pas les ensembles M_1 autres que M' qui diffèrent de M par un seul élément; on exclura de même les ensembles M_2 autres que M'', et ainsi de suite; les choix faits dans les M_1 (autres que M'), dans les M_2 (autres que M'') sont ainsi superflus; on peut les éviter et l'on est ainsi ramené à la première démonstration.

A un autre point de vue, on peut se demander s'il est aisé de donner une règle de choix plus simple que celle qui résulterait d'une ordination préalable de l'ensemble. Je me suis posé la question pour l'ensemble M très simple formé des nombres rationnels compris entre 0 et 1; les divers moyens que j'ai pu imaginer pour attacher à tout sous-ensemble A de M un élément distingué a, reviennent tous à supposer établie une règle pour bien ordonner M et à attacher à chaque A son premier élément a ou un élément dont le calcul exige la connaissance préalable de ce premier élément. Ce sont là du moins les moyens qui se présentent naturellement. Si, partant d'une ordination déterminée de M, on distingue entre les ensembles A qui font partie de la plus petite chaîne C et les ensembles J qui n'en font pas partie, il est possible de compliquer les règles en attribuant à chacun de ces ensembles J un élément distingué autre que son premier élément: mais cette complication artificielle exige encore davantage la connaissance de l'ordination préalable.

Le « théorème » de M. Zermelo revient donc à ceci : on peut remplacer des choix successifs par des choix simultanés transfiniment plus nombreux, et choisir ensuite entre ces choix. Par exemple, à défaut d'une règle, qui permettrait de calculer patiemment, les uns après les autres, les chiffres décimaux d'un nombre tel que $\sqrt{2}$ ou que π, on pourrait écrire *tous* les nombres décimaux et se proposer, les embrassant d'un seul regard, d'après une infinité de conditions imposées, de choisir parmi eux celui qui est égal à $\sqrt{2}$ ou à π.

IV. — Cinq lettres sur la théorie des ensembles (1).

I. — *Lettre de M. Hadamard à M. Borel.*

J'ai lu avec intérêt les arguments que tu opposes (2e Cahier du Tome LX des *Mathematische Annalen*) à la démonstration de M. Zermelo parue dans le Tome précédent. Je ne partage cependant pas ton opinion à ce sujet. Je n'admets pas, tout d'abord, l'assimilation que tu établis entre le fait qui sert de point de départ à M. Zermelo et le raisonnement qui consisterait à numéroter les éléments de l'ensemble les uns après les autres, ce numérotage étant poursuivi *transfiniment*. Il y a, en effet, une différence fondamentale entre les deux cas : le raisonnement qui vient d'être

(1) *Bulletin de la Société mathématique de France*, décembre 1904.

cité en dernier lieu comporte une série de choix successifs *dont chacun dépend des précédents;* c'est pour cela que son application transfinie est inadmissible. Je ne vois aucune analogie à établir, au point de vue qui nous occupe, entre les choix en question et ceux dont parle M. Zermelo, lesquels sont *indépendants les uns des autres.*

C'est d'ailleurs dans le cas d'une infinité *non dénombrable* de choix que tu récuses cette manière d'opérer; mais, à mon tour, je ne vois pas de différence, à cet égard, entre le cas d'une infinité non dénombrable et celui d'une infinité dénombrable. La différence serait manifeste s'il y avait une dépendance quelconque entre les choix en question, parce qu'il faudrait alors avoir égard à l'ordre dans lequel on les opérerait : elle me paraît, encore une fois, s'évanouir complètement dans le cas des choix indépendants.

Ce qui est certain, c'est que M. Zermelo ne donne aucun moyen d'exécuter *effectivement* l'opération dont il parle, et qu'il reste douteux que personne puisse, dans la suite, indiquer ce moyen. Il aurait été assurément plus intéressant de résoudre le problème sous cette forme; mais la question ainsi posée (détermination effective de la correspondance cherchée) n'en est pas moins complètement distincte de celle que nous examinons (une telle correspondance existe-t-elle?) : il y a entre elles toute la différence, laquelle est fondamentale, qui existe entre ce que M. Tannery (¹) appelle *une correspondance* qui peut être *définie* et une correspondance qui peut être *décrite.* Plusieurs questions importantes de Mathématiques changeraient totalement de sens, et de solutions, si l'on substituait le second mot au premier. Tu emploies des correspondances dont tu constates l'*existence* sans pouvoir cependant les *décrire*, dans ton important raisonnement relatif aux séries qui admettent leur cercle de convergence comme coupure : si l'on se bornait aux séries entières dont la loi de formation peut être décrite, l'opinion ancienne (à savoir, que les séries entières admettant leur cercle de convergence comme coupure sont l'exception) devrait, à mon sens, être considérée comme la vraie. C'est d'ailleurs une pure question de sentiment; car la notion de correspondance « qui peut être décrite » est, pour reprendre ton expression, « en dehors des Mathématiques »; elle relève du domaine de la psychologie et est relative à une propriété de notre esprit; c'est une question de cette nature que celle de savoir si la correspondance employée par M. Zermelo pourra jamais être indiquée *en fait.*

Quant à l'existence de cette correspondance, elle me paraît aussi adéquate à la possibilité de prendre *un* élément dans un ensemble quelconque donné, que la proposition suivante :

A. *Un nombre x étant donné, il existe des nombres y qui ne sont liés à x par aucune équation algébrique à coefficients entiers,*

l'est à celle-ci :

(¹) *Revue générale des Sciences*, t. VIII, 1897, p. 133 et suiv.

B. *Il existe des fonctions y de x telles que, pour aucune valeur de x, y n'ait ni valeur algébrique, ni une valeur liée à x par une équation algébrique à coefficients entiers.*

On pourrait d'ailleurs, sans doute, former de telles fonctions. Mais ce que je prétends, c'est que cela n'est nullement nécessaire pour affirmer l'exactitude du théorème B; et je crois que beaucoup de mathématiciens ne prendraient pas plus que moi cette peine s'ils avaient à employer le théorème en question. J. HADAMARD.

II. — *Lettre de M. Baire à M. Hadamard.*

Borel me communique la lettre où vous lui exposez votre manière de voir sur le grand débat soulevé par la Note Zermelo. Je vous demande la permission de vous adresser quelques réflexions qu'elle me suggère.

Je suis, vous le savez, de l'avis de Borel, en gros, et si je m'en écarte, ce sera pour aller plus loin que lui.

Supposons qu'on fasse un effort pour essayer d'appliquer la méthode de Zermelo à l'ensemble M des suites d'entiers positifs. On prend dans M un élément distingué m_1; reste l'ensemble M — m_1, dans lequel on prend un élément distingué m_2; etc. Ces choix successifs dépendent bien chacun de ceux qui les précèdent. Mais, dites-vous avec M. Zermelo, les choix sont indépendants les uns des autres, parce qu'il admet comme point de départ *un choix d'élément distingué fait dans* TOUTE *partie de* M. Ceci ne me paraît pas satisfaisant : c'est, pour moi, dissimuler la difficulté *en la noyant dans une difficulté plus grande.*

L'expression *ensemble donné* est employée à chaque instant : a-t-elle un sens? Pas toujours, selon moi. Dès qu'on parle d'infini (même dénombrable, et c'est ici que je suis tenté d'être plus radical que Borel), l'assimilation, *consciente ou inconsciente*, avec un sac de billes qu'on donne de la main à la main, doit complètement disparaître, et nous sommes, à mon avis, dans le *virtuel*, c'est-à-dire que nous faisons des conventions qui nous permettent ultérieurement, un objet étant défini *par une nouvelle convention*, d'affirmer certaines propriétés de cet objet. Mais croire qu'on est allé plus loin ne me paraît pas légitime. En particulier, de ce qu'un ensemble est donné (nous serons d'accord pour dire, par exemple, que nous nous donnons l'ensemble des suites d'entiers positifs), *il est faux pour moi de considérer les parties de cet ensemble comme données.* A plus forte raison je refuse d'attacher un sens au fait de concevoir un choix fait dans chaque partie d'un ensemble.

M. Zermelo dit : « Concevons qu'à tout ensemble partiel de M corresponde un de ses éléments. » C'est là une conception qui n'a rien de contradictoire, d'accord. Aussi, tout ce qu'il démontre pour moi, c'est que nous n'apercevons pas de contradiction à concevoir que, dans tout ensemble qu'on nous définira, les éléments aient entre eux des relations de position identiques à celles qu'ont les éléments des ensembles bien ordon-

nés. Pour dire après cela que tout ensemble peut être mis sous la forme d'un ensemble bien ordonné, il faut donner aux mots une extension extraordinaire et, j'ajouterai, trompeuse.

De ce qui précède, je ne suis arrivé que bien incomplètement à rendre ma pensée. J'ai dit ma manière de voir dans la phrase qu'a bien voulu transcrire Borel dans sa Note. Pour moi, le progrès, dans cet ordre d'idées, consisterait à délimiter le domaine de ce qui est définissable. Et, en fin de compte, en dépit des apparences, tout doit se ramener au fini.

Décembre 1904. R. Baire.

III. — *Lettre de M. Lebesgue à M. Borel.*

Vous me demandez mon opinion sur la Note de M. Zermelo (*Math. Annalen*, t. LIX), sur les objections que vous lui avez faites (*Math. Annalen*, t. LX) et sur la lettre de M. Hadamard que vous me communiquez; la voici. Excusez-moi d'être long, j'ai essayé d'être clair.

Tout d'abord je suis d'accord avec vous pour ceci : M. Zermelo a très ingénieusement démontré qu'on savait résoudre le problème A :

A. *Mettre un ensemble* M *sous forme bien ordonnée*,

toutes les fois qu'on savait résoudre le problème B :

B. *Faire correspondre à chaque ensemble* M′ *formé avec des éléments de* M *un élément particulier* m′ *de* M′.

Malheureusement, le problème B n'est facile à résoudre, à ce qu'il semble, que pour les ensembles qu'on sait bien ordonner; par suite, on n'a pas une solution générale du problème A.

Je doute fort qu'on puisse donner une solution générale de ce problème, du moins si l'on admet, avec M. Cantor, que définir un ensemble M c'est nommer une propriété P appartenant à certains éléments d'un ensemble N précédemment défini et caractérisant, par définition, les éléments de M. En effet, avec cette définition, on ne sait rien sur les éléments de M d'autre que ceci : ils possèdent tous les propriétés *inconnues* des éléments de N et ce sont les seuls qui ont la propriété P *inconnue*. Rien là-dedans ne permet de distinguer deux éléments de M, encore moins de les classer comme il faudrait le faire pour résoudre A.

Cette objection, faite *a priori* à tout essai de solution de A, tombe évidemment si l'on particularise N ou P; l'objection tombe, par exemple, si N est l'ensemble des nombres. Tout ce qu'on peut espérer faire de général, c'est indiquer des problèmes, tels que B, dont la résolution entraînerait celle de A et possibles dans certains cas particuliers, mais qui se rencontrent fréquemment. D'où l'intérêt, à mon avis, du raisonnement de M. Zermelo.

Je crois que M. Hadamard est plus fidèle que vous à la pensée de M. Zermelo en interprétant la Note de cet auteur comme un essai, non pas de

résolution effective de A, mais de démonstration d'existence de la solution. La question revient à celle-ci, peu nouvelle : *Peut-on démontrer l'existence d'un être mathématique sans le définir?*

C'est évidemment une affaire de convention ; mais je crois qu'on ne peut bâtir solidement qu'*en admettant qu'on ne démontre l'existence d'un être qu'en le définissant*. A ce point de vue, voisin de celui de Kronecker et de M. Drach, il n'y a pas à distinguer entre A et le problème C :

C. *Tout ensemble peut-il être bien ordonné?*

Je n'aurais rien de plus à dire si la convention que j'ai indiquée était universellement admise ; mais je dois avouer qu'on emploie souvent, et que j'ai moi-même souvent employé, le mot *existence* dans d'autres sens. Par exemple, lorsqu'on interprète un raisonnement bien connu de M. Cantor en disant : *il existe une infinité non dénombrable de nombres*, on ne donne cependant pas le moyen de nommer une telle infinité. On montre seulement, vous l'avez dit avant moi, que, chaque fois qu'on aura une infinité dénombrable de nombres, on pourra définir un nombre ne faisant pas partie de cette infinité. (Le mot *définir* a tout le temps le sens de : *nommer une propriété caractéristique du défini.*) Une existence de cette nature peut être utilisée dans un raisonnement et de la manière suivante : une propriété est vraie, si, la nier, conduit à admettre qu'on peut ranger tous les nombres en suite dénombrable. Je crois qu'elle ne peut intervenir que de cette manière.

M. Zermelo utilise l'*existence* d'une *correspondance* entre les sous-ensembles de M et certains de leurs éléments. Vous voyez que, quand même l'existence de ces correspondances serait hors de doute, suivant la manière dont cette existence aurait été prouvée, il ne serait pas évident qu'on ait le droit d'utiliser cette existence comme le fait M. Zermelo.

J'arrive au raisonnement que vous énoncez ainsi : « Il est possible, dans un ensemble particulier M', de choisir *ad libitum* l'élément distingué *m'*; ce choix pouvant être fait pour chacun des ensembles M' peut être fait pour l'ensemble de ces ensembles », et duquel semble résulter l'existence des correspondances.

Tout d'abord, M' étant donné, est-il évident qu'on puisse choisir *m'*? Cela serait évident si M' existait, au sens presque kroneckérien que j'ai dit, puisque dire que M' existe serait alors affirmer qu'on sait nommer certains de ses éléments. Mais étendons le sens du mot *exister*. L'ensemble Γ des correspondances entre les sous-ensembles M' et les éléments distingués *m'* *existe* certainement pour MM. Hadamard et Zermelo ; ce dernier représente même le nombre de ses éléments par un produit transfini. Cependant, sait-on choisir un élément de Γ? Non, évidemment, puisque ce serait donner de B, pour M, une solution déterminée.

Il est vrai que j'emploie le mot *choisir* dans le sens de *nommer* et qu'il suffit peut-être pour le raisonnement de M. Zermelo que *choisir* signifie *penser à*. Mais il faut cependant remarquer qu'on n'indique pas celui auquel on pense et qu'il est néanmoins nécessaire au raisonnement de

M. Zermelo qu'on pense à *une correspondance déterminée toujours la même.* M. Hadamard croit, il me semble, qu'il n'est pas nécessaire qu'on démontre qu'on peut *déterminer* un élément (et un seul); c'est de là, à mon avis, que viennent les différences d'appréciation.

Pour mieux vous faire sentir la difficulté que je vois, je vous rappelle que, dans ma thèse, j'ai démontré l'existence (sens non kroneckérien et peut-être difficile à préciser) d'ensembles mesurables non mesurables B, mais il restait douteux pour moi qu'on pût jamais en nommer un. Dans ces conditions, aurais-je eu le droit de fonder un raisonnement sur cette hypothèse : *je suppose choisi un ensemble mesurable non mesurable* B, alors que je doutais que personne pût jamais en nommer un ?

Ainsi je vois déjà une difficulté dans ceci « dans un M′ déterminé je puis choisir un m' déterminé », puisqu'il existe des ensembles (l'ensemble C par exemple, qu'on pourrait considérer comme un ensemble M′ provenant d'un ensemble plus général) dans lesquels il est peut-être impossible de choisir un élément. Il y a ensuite la difficulté que vous signalez relative à l'infinité des choix, ce qui fait que, si l'on veut considérer le raisonnement de M. Zermelo comme tout à fait général, il faut admettre qu'on parle d'une infinité de choix, infinité de puissance peut-être grande; on ne donne d'ailleurs ni la loi de cette infinité, ni la loi d'un des choix; on ne sait pas s'il est possible de nommer une loi définissant un ensemble de choix ayant la puissance de l'ensemble des M′; on ne sait pas s'il est possible, étant donné un M′, de nommer un m'.

En résumé, quand j'examine de près le raisonnement de M. Zermelo, comme d'ailleurs plusieurs raisonnements généraux sur les ensembles, je le trouve trop peu kroneckérien pour lui attribuer un sens (en tant que théorème d'existence de la solution de C, seulement, bien entendu).

Vous faites allusion à ce raisonnement : « Pour bien ordonner un ensemble il suffit d'y choisir un élément, puis un autre, etc. » Il est certain que ce raisonnement présente des difficultés énormes, plus grandes encore, au moins en apparence, que celui de M. Zermelo; et je suis tenté de croire avec M. Hadamard qu'il y a progrès à avoir remplacé une infinité de choix successifs et dépendant les uns des autres par une infinité, non ordonnée, de choix indépendants. Il n'y a peut-être là qu'une illusion et la simplification apparente tient peut-être seulement à ce qu'on doit remplacer une infinité ordonnée de choix par une infinité non ordonnée, mais de puissance plus grande. De sorte que le fait qu'on peut ramener à la seule difficulté, placée au début du raisonnement de M. Zermelo, toutes les difficultés du raisonnement simpliste que vous citez prouve peut-être simplement que cette seule difficulté est très grande. En tout cas, elle ne me paraît pas disparaître parce qu'il s'agit d'un ensemble non ordonné de choix indépendants. Par exemple, si je crois à l'existence de fonctions $y(x)$ telles que, quel que soit x, y ne soit jamais lié à x par une équation algébrique à coefficients entiers, c'est parce que je crois, avec M. Hadamard, qu'il est possible d'en construire; mais ce n'est pas, pour moi, la consé-

quence immédiate de l'existence, quel que soit x, de nombres y qui ne soient liés à x par aucune équation à coefficients entiers ([1]).

Je suis pleinement d'accord avec M. Hadamard quand il déclare que la difficulté qu'il y a à parler d'une infinité de choix sans en donner la loi est aussi grave, qu'il s'agisse ou non d'une infinité dénombrable. Quand on dit, comme dans le raisonnement que vous critiquiez, « ce choix pouvant être fait pour chacun des ensembles M′, peut être fait pour l'ensemble de ces ensembles », on ne dit rien si l'on n'explique pas les termes employés. Faire un choix, ce peut être écrire ou nommer l'élément choisi ; faire une infinité de choix, ce ne peut être écrire ou nommer les éléments choisis, un à un : la vie est trop courte. Il faut donc dire ce que c'est faire. On entend par là, en général, que c'est donner la loi qui définit les éléments choisis, mais cette loi est pour moi, comme pour M. Hadamard, aussi indispensable, qu'il s'agisse d'une infinité dénombrable ou non.

Peut-être cependant suis-je encore d'accord avec vous sur ce point parce que, si je n'établis pas de différences théoriques entre les deux infinités, au point de vue pratique, je fais une grande différence entre elles. Quand j'entends parler d'une loi définissant une infinité transfinie de choix, je suis très méfiant, parce que je n'ai jamais encore vu de pareilles lois, tandis que je connais des lois définissant une infinité dénombrable de choix. Mais ce n'est qu'une affaire de routine et, à la réflexion, je vois parfois des difficultés aussi graves, à mon avis, dans des raisonnements où n'interviennent qu'une infinité dénombrable de choix que dans des raisonnements où il y en a une transfinité. Par exemple, si je ne considère pas comme établi par le raisonnement classique que tout ensemble de puissance supérieure au dénombrable contient un ensemble dont la puissance est celle de l'ensemble des nombres transfinis de la classe II de M. Cantor, je n'attribue pas plus de valeur à la méthode par laquelle on démontre qu'un ensemble non fini contient un ensemble dénombrable. Bien que je doute fort qu'on nomme jamais un ensemble qui ne soit ni fini, ni infini, l'impossibilité d'un tel ensemble ne me paraît pas démontrée. Mais je vous ai déjà parlé de ces questions.

H. Lebesgue.

IV. — *Lettre de M. Hadamard à M. Borel.*

La question me paraît tout à fait claire maintenant, après la lettre de M. Lebesgue. De plus en plus nettement, elle tient tout entière dans la distinction, exposée dans l'article de M. Tannery, entre ce qui est *déterminé* et ce qui est *décrit*.

([1]) En corrigeant les épreuves, j'ajoute qu'en fait le raisonnement, par lequel on légitime ordinairement l'énoncé A de M. Hadamard (p. 151), légitime en même temps l'énoncé B (p. 152). Et, à mon avis, c'est parce qu'il légitime B qu'il légitime A.

Lebesgue, Baire et toi, adoptez à cet égard la manière de voir de Kronecker, que je croyais jusqu'ici lui être particulière. Vous répondez négativement à la question posée (ci-dessus, p. 154) par M. Lebesgue : Peut-on démontrer l'existence d'un être mathématique sans le définir? J'y réponds affirmativement. Je prends pour mienne, autrement dit, la réponse que Lebesgue fait lui-même à son objection relative à l'ensemble Γ.

Qu'il *nous* soit impossible, au moins actuellement, de *nommer* un élément de cet ensemble, j'en conviens. C'est là la question pour vous; ce ne l'est pas pour moi.

Il n'y a qu'un point sur lequel il me semble que Lebesgue ne soit pas logique avec lui-même. C'est lorsqu'il se reconnaît ou ne se reconnaît pas le droit d'utiliser une existence, suivant la manière doni elle a été démontrée. Pour moi, les *existences* dont il parle sont des faits comme les autres. Sinon, elles n'ont pas lieu.

La question se pose de même vis-à-vis de Baire. Je n'aimerais pas beaucoup la placer, comme il le fait (p. 152), à la façon de M. Hilbert, sur le terrain du *non contradictoire*, qui me paraît encore relever de la psychologie et faire entrer en ligne de compte les propriétés de nos cerveaux. Je ne comprends même pas bien comment M. Zermelo peut avoir *démontré* que *nous n'apercevons* pas de contradiction, etc. Cela ne se *démontre* pas, cela se *constate* : on en a aperçu ou l'on n'en a pas aperçu.

Ce point écarté, la question principale, celle de savoir si l'ensemble peut être ordonné, n'a évidemment pas pour Baire (pas plus que pour Lebesgue et toi) le même sens que pour moi. Je dirais plutôt : l'ordination est-elle possible? (et non pas même peut-*on* ordonner, de crainte d'avoir à penser à ce qu'est cet *on*) : Baire dirait : pouvons-*nous* ordonner? Question toute subjective, à mon avis.

Ce sont donc deux conceptions des Mathématiques, deux mentalités qui sont en présence. Je ne vois, dans tout ce qui a été dit jusqu'ici, aucun motif de changer la mienne. Je ne prétends pas l'imposer. Tout au plus ferai-je valoir en sa faveur les arguments que j'ai indiqués dans la *Revue générale des Sciences* (30 mars 1905), savoir :

1° Je crois que le débat est au fond le même qui s'est élevé entre Riemann et ses prédécesseurs, sur la notion même de fonction. La *loi* qu'exige Lebesgue me paraît ressembler fort à l'expression [1] analytique que réclamaient à toute force les adversaires de Riemann. Et même à une expression analytique pas trop bizarre. Non seulement la *numérabilité* des choix

(1) Je crois devoir insister un peu sur ce point de vue qui, s'il faut dire toute ma pensée, me paraît former le fond même du débat. Il me semble que le progrès véritablement essentiel des Mathématiques, à partir de l'invention même du Calcul infinitésimal, a consisté dans l'annexion de notions successives qui, les unes pour les Grecs, les autres pour les géomètres de la Renaissance ou les prédécesseurs de Riemann, étaient « en dehors des Mathématiques », parce qu'il était impossible de les décrire.

ne me paraît pas changer la question, mais il en est de même de l'*unicité*. Je ne vois pas comment nous aurions le droit de dire : « Pour chaque valeur de x il existe un nombre satisfaisant à .... Soit y ce nombre ... », alors que, parce que « la mariée est trop belle », nous ne pouvons pas dire : « Pour chaque valeur de x il existe une infinité de nombres satisfaisant à .... Soit y l'un de ces nombres ... ».

2° Les choix arbitraires de Tannery conduisent à des nombres ν, que *nous serions* incapables de définir. Je ne conçois pas que ces nombres n'existent pas.

Quant aux raisonnements présentés par M. Bernstein (*Math. Annalen*, t. IX, p. 187), et, par conséquent, à ses objections à la démonstration de M. Zermelo, je ne les considérerais pas, pour ma part, comme probants. Cette opinion est d'ailleurs indépendante de la question que nous discutons actuellement.

M. Bernstein part du paradoxe de M. Burali-Forti (*Circolo matematico di Palermo*, 1897) relatif à l'ensemble W de *tous* les nombres ordinaux. Pour échapper à la contradiction mise en évidence par M. Burali-Forti, il suppose le nombre ordinal W tel qu'il soit impossible de lui ajouter 1. Cette opinion est, pour moi, inadmissible, ainsi que les arguments imaginés en sa faveur par M. Bernstein. L'ordre établi (d'après la théorie de M. Cantor) entre les éléments de W et l'élément supplémentaire (c'est à cet ordre que s'attaque l'auteur) est une pure *convention*, qu'on est toujours libre de faire et à laquelle les propriétés de W, quelles qu'elles soient, ne sauraient mettre aucun obstacle.

La solution est autre. C'est l'existence même de l'ensemble W qui implique contradiction. Dans sa définition, la définition générale du mot *ensemble* est incorrectement appliquée. On n'a le droit de former un ensemble qu'avec des objets préalablement existants et il est aisé de voir que la définition de W suppose le contraire.

Même observation pour *l'ensemble de tous les ensembles* (Hilbert, Congrès de Heidelberg).

Revenons à la question primitive. Voici encore, à cet égard, non un argument, car je crois que nous coucherons éternellement sur nos positions, mais une conséquence de tes principes.

Cantor a considéré l'ensemble de toutes les fonctions qui, dans l'intervalle (0, 1), ne prennent que les valeurs 0, 1. Cet ensemble a, pour moi, un sens clair et sa puissance est $2^{\aleph}$, comme l'énonce Cantor. De même, l'ensemble de toutes les fonctions de x a pour moi un sens, et je vois clairement que sa puissance est $\aleph^{\aleph}$.

Quel sens tout cela a-t-il pour toi? Il me paraît évident que cela ne peut en avoir aucun. Car à toute fonction tu imposes une condition supplémentaire qui n'a aucun sens mathématique : celle d'être *descriptible* pour *nous*.

Ou plutôt, voici ce que cela signifie : on ne doit considérer, à ton point de vue, que les fonctions définissables en un nombre fini de mots. Mais, à

ce compte, les deux ensembles ainsi formés sont *dénombrables*, ainsi que tous les ensembles possibles, d'ailleurs.

J. HADAMARD.

V. — *Lettre de M. Borel à M. Hadamard.*

... Je voudrais d'abord te signaler une intéressante remarque faite par M. Lebesgue à la séance de la Société du 4 mai : Comment M. Zermelo peut-il être assuré qu'aux divers points de son raisonnement il parle *du même* choix de l'élément distingué, puisqu'il ne le caractérise par rien *pour lui-même* (il ne s'agit même pas ici d'un contradicteur possible; il s'agit d'être cohérent avec soi-même).

Quant à ta nouvelle objection, voici quelle est ma situation à son égard. Je n'aime guère écrire des alephs, mais je consens cependant à faire des raisonnements équivalents à ceux dont tu parles, sans me faire guère illusion sur leur valeur intrinsèque, *mais en les regardant comme pouvant guider pour d'autres raisonnements plus sérieux.* Comme exemple pratique, je puis te citer la Note III que j'ai insérée à la fin de mon dernier petit Livre (*Leçons sur les fonctions de variables réelles, etc.*, rédigées par Maurice Fréchet); le raisonnement qui y est employé est manifestement suggéré par le raisonnement de M. Cantor, que j'ai rapporté dans mes premières *Leçons sur la théorie des fonctions* (1), page 107.

La forme que j'adopte dans cette Note III n'est pas encore absolument satisfaisante, comme je l'indique au bas de la dernière page de mon Livre; mais le raisonnement analogue de M. Lebesgue dans son Mémoire paru dans le *Journal de Jordan* (1905) est, je crois, tout à fait irréprochable, en ce sens qu'il conduit à un résultat précis, exprimable au moyen d'un nombre fini de mots; il a cependant son origine dans celui de M. Cantor.

On peut se demander quelle est la valeur réelle de ces raisonnements que je ne regarde pas comme valables absolument et qui cependant conduisent ultérieurement à des résultats effectifs. Il semble en effet que, s'ils étaient dépourvus de toute valeur, ils ne pourraient conduire à rien, car ce seraient des assemblages de mots vides de sens. Je crois qu'on serait ainsi trop sévère et qu'ils ont une valeur analogue à celle de certaines théories de Physique mathématique, par lesquelles nous ne prétendons pas exprimer la réalité, mais avoir un guide qui nous permette, par analogie, de découvrir des phénomènes nouveaux, qu'il reste ensuite à vérifier. Il y aurait un travail considérable à faire pour savoir quel est le sens réel et précis qu'on peut attribuer à des raisonnements de ce genre; ce travail est inutile ou du moins hors de proportion avec son utilité; les rapports avec le concret de ces raisonnements trop abstraits apparaissent d'eux-mêmes lorsque le besoin s'en fait sentir.

---

(1) Dans les Notes I et II de ce petit Livre, je fais constamment des raisonnements du type de ceux que tu me refuses le droit de faire; je suis d'ailleurs à chaque instant rempli de scrupules et chacune de ces deux Notes se termine par une phrase très restrictive.

Je serai d'accord avec toi sur le fait qu'il est contradictoire de parler de l'ensemble de tous les ensembles, car, par le raisonnement de la page 107 citée plus haut, on peut former un ensemble de puissance plus grande, mais je crois que cette contradiction tient à ce qu'on introduit des ensembles non définis réellement.

Em. Borel.

V. — Sur les principes de la théorie des ensembles (1).

Ces principes ont été beaucoup discutés en ces dernières années; je voudrais indiquer brièvement les conclusions personnelles auxquelles je suis arrivé à ce sujet, sans entrer dans la discussion de tous les arguments émis.

L'étude de ces arguments n'a fait que rendre plus nette l'opinion que j'émettais sous forme atténuée, il y a dix ans, dans mes premières *Lecons sur la théorie des fonctions :* « Nos connaissances précises sur les puissances diverses n'excèdent *guère* la remarque suivante : il y a des ensembles dénombrables et des ensembles non dénombrables, cette dernière notion étant *surtout* négative ». J'écrirais aujourd'hui « n'excèdent *pas* la remarque suivante » et « cette notion étant *purement* négative ». Il ne me semble pas, en effet, que les efforts tentés dans ces dernières années aient résolu d'une manière satisfaisante ce que j'ai appelé l' « antinomie du transfini ». Je ne m'égarerai pas en discussions métaphysiques sur le sens du mot « indéfiniment »; que l'emploi de ce mot soulève des difficultés pour les philosophes, c'est un fait sans importance pour les mathématiciens : il leur suffit de savoir qu'ils s'entendent parfaitement entre eux, sans craindre aucune ambiguïté. Lorsqu'un de nous dit qu'il considère la suite naturelle des nombres entiers, chacun comprend, et est assuré de comprendre *la même chose* que son voisin; c'est évidemment là le seul criterium possible de la validité d'un langage, celui auquel on est toujours forcé de revenir. Car les prétendus systèmes entièrement logiques reposent toujours sur le postulat de l'existence de la langue vulgaire; ce langage commun à des millions d'hommes, et avec lequel ils s'entendent à peu près entre eux, nous est donné comme un fait, qui impliquerait un grand nombre de cercles vicieux, s'il fallait le créer *ex nihilo*.

J'ai indiqué, il y a quelques années (2), que l'antinomie du transfini se résoudrait si les mathématiciens arrivaient à avoir une idée commune du mot *transfiniment* aussi claire que celle qu'ils ont du mot *indéfiniment*. En posant ainsi la question, je ne prétendais pas la résoudre; je voulais surtout attirer l'attention sur le fait que notre conception de l'*indéfini*

---

(1) *IVe Congrès international des Mathématiciens*, Rome 1908.

(2) *A propos de l'infini nouveau*, 1899; *l'antinomie du transfini*, 1900 (*Revue philosophique*). (§ I et II de cette Note IV.)

renferme quelque chose de plus que la proposition : *après chaque entier il y en a un autre;* car la proposition analogue : *au delà de chaque suite indéfinie de fonctions croissantes, il y en a une autre*, ne suffit pas à nous donner une idée nette du transfini.

Je serais aujourd'hui disposé, à la suite des efforts faits dans ces dernières années pour donner un sens clair au *transfini*, à résoudre la question par la négative, pour autant qu'il est possible d'affirmer un fait aussi purement négatif; je crois qu'il n'est pas possible d'arriver à une définition du transfini nous apparaissant comme aussi claire que celle de l'indéfini. Je pourrais me borner à constater les désaccords du grand nombre d'esprits éminents qui ont approfondi ces questions; mais je voudrais essayer de fournir un argument plus positif, qui n'est pas sans analogie avec les « paradoxes » connus, mais qui me paraît cependant échapper aux principales objections qu'on leur a faites.

Il est bien connu que tout système de conventions énoncé en un nombre limité de mots, parmi lesquels peut figurer le mot *indéfiniment*, ne conduira jamais qu'à noter un ensemble dénombrable; il n'est donc pas possible qu'on arrive à une notation bien définie pour l'ensemble des nombres que M. Cantor appelle *nombres de la seconde classe*. Lorsqu'on raisonne sur ces nombres, ou sur un ensemble de même puissance, on ne peut donc faire que des raisonnements généraux ou symboliques, dans lesquels l'ensemble considéré est représenté par un symbole unique. De tels raisonnements, en tant que raisonnements généraux, sont légitimes du moment qu'ils sont exempts de contradiction, mais ils sont en même temps vides de tout contenu précis. Pour pouvoir leur donner un contenu, il faudrait préciser la désignation des éléments de l'ensemble, et c'est précisément ce qui est impossible. La théorie des ensembles non dénombrables se réduit donc forcément à une sorte d'algèbre logique dont les symboles ne recouvrent aucune réalité accessible, les divers mathématiciens ne pouvant être assurés qu'ils sont d'accord sur cette réalité, puisqu'ils n'en ont pas une représentation commune.

De tels raisonnements généraux et abstraits ne sont cependant pas toujours inutiles, car ils contribuent parfois à éclaircir les idées et peuvent, de plus, être souvent transposés dans un domaine plus concret; mais on ne doit pas se faire d'illusion sur le contenu réel de ces raisonnements généraux.

Il est nécessaire de dire un mot de la notion du *continu*, qui est le seul exemple d'ensemble non dénombrable qui soit bien connu, c'est-à-dire duquel les mathématiciens aient une notion claire et commune (ou croient avoir, ce qui est pratiquement la même chose). Je regarde cette notion comme acquise par l'intuition géométrique; on sait que la notion arithmétique complète du continu exige qu'on admette la légitimité d'une infinité dénombrable de choix successifs et arbitraires. Cette légitimité me paraît fort discutable, mais on doit cependant la distinguer essentiellement de la légitimité d'une infinité non dénombrable de choix (successifs ou simultanés). Cette dernière notion me paraît, comme j'ai déjà eu l'occasion de

le dire, entièrement vide de sens. Lorsqu'il s'agit d'une infinité dénombrable de choix, on ne peut évidemment pas les effectuer tous, mais on peut du moins indiquer une marche telle ue, cette marche étant fixée d'avance, on soit assuré que l'un quelconque des choix sera effectué au bout d'un temps fini; si donc deux systèmes donnés de choix sont distincts, on est certain qu'on s'en apercevra au bout d'un nombre fini d'opérations. Lorsque l'infinité des choix n'est pas dénombrable, il n'est pas possible de concevoir un moyen de la définir, c'est-à-dire de la distinguer d'une infinité analogue : il n'est donc pas possible de la regarder comme un être mathématique, pouvant être introduit dans les raisonnements.

Je devrais, pour terminer ce bref exposé, expliquer brièvement comme je concilie la notion du continu avec mon opinion sur notre défaut de notion des ensembles non dénombrables. Le continu ne m'apparaît jamais comme donné dans son intégralité, au point de vue arithmétique; chacun de ses éléments peut être défini (ou du moins, il n'est aucun de ses éléments dont nous puissions actuellement affirmer qu'il ne peut pas être défini) (1); mais l'ensemble des éléments effectivement définis sera toujours dénombrable; c'est seulement sous une forme négative que nous sommes assurés de ne jamais l'épuiser ainsi.

VI. — Les « Paradoxes » de la théorie des ensembles (2).

Nous prendrons comme type de ces paradoxes le raisonnement bien connu de M. J. Richard relatif aux nombres (fractions décimales) qui peuvent être définis au moyen d'un nombre fini de mots. L'ensemble de ces nombres est dénombrable et de plus peut être ordonné : il suffit de ranger les définitions d'après le nombre de lettres qu'exige leur énoncé, celles qui emploient le même nombre de lettres étant classées alphabétiquement, comme les mots dans un dictionnaire. Mais, d'autre part, étant donné un ensemble dénombrable ordonné de fractions décimales,

$$(1) \qquad \alpha_1, \quad \alpha_2, \quad \ldots, \quad \alpha_n, \quad \ldots,$$

il est aisé de définir une fraction α bien déterminée et distincte de chacune d'elles; il suffit de supposer que le $n^{\text{ième}}$ chiffre décimal de α se déduit suivant une loi déterminée du $n^{\text{ième}}$ chiffre décimal de α_n, cette loi

(1) Voici, pour donner une idée de mon point de vue, un problème qui me paraît être des plus importants dans la théorie arithmétique du continu : Est-il ou non possible de définir un ensemble E tel qu'on ne puisse nommer aucun élément individuel de cet ensemble E, c'est-à-dire le distinguer sans ambiguïté de tous les autres éléments de E. [*Voir* LEBESGUE, *Sur les fonctions représentables analytiquement* (*Journal de Jordan*, 1905).]

(2) *Annales de l'École Normale*, octobre 1908.

étant telle que ces deux chiffres soient sûrement distincts; on peut, par exemple, aux chiffres 0, 1, 2, 3, ..., 9, faire correspondre respectivement 9, 8, 7, ..., 0, ou bien 1, 2, 3, ..., 9, 0, etc. La fraction α est bien définie au moyen d'un nombre fini de mots (puisque sa définition complète vient d'être donnée) et cependant elle n'appartient pas à la suite (1); il y a donc contradiction.

Les diverses explications données de cette contradiction ne me paraissant pas satisfaisantes ([1]), je voudrais reprendre la question en me plaçant sur le terrain des réalités, sans y mêler aucune considération métaphysique ni de logique pure. Pour cela, je préciserai tout d'abord que *je considère les nombres décimaux qui sont définis d'une manière précise et sans ambiguïté possible au moyen d'un nombre fini de mots.* Si, en effet, à une définition correspondait une infinité de nombres, l'ensemble de ces nombres, à supposer même qu'il fût dénombrable, ne se rangerait pas dans une suite telle que (1) et le nombre α ne pourrait donc être déterminé.

Ainsi, je considère les suites de lettres présentant, en langue française, un sens bien déterminé et définissant d'une manière précise un nombre unique. Pour former toutes ces suites, on pourra procéder comme il suit : on considérera un nombre déterminé N de caractères, choisis parmi les quarante lettres de l'alphabet ou signes divers de ponctuation; leurs combinaisons possibles sont au nombre de N^{40} parmi lesquelles, bien évidemment, la plupart n'ont aucun sens; on conservera uniquement celles de ces combinaisons qui, non seulement ont un sens mathématique, mais définissent un nombre unique et bien déterminé; l'ensemble de ces nombres est sûrement dénombrable, puisque les définitions le sont; je dis que *cet ensemble* E *renferme tout nombre α qui peut, au moyen d'un nombre fini de mots, être défini sans ambiguïté possible*, ce qui signifie que deux mathématiciens quelconques, à chacun desquels on communiquerait la définition, construiront le même nombre, c'est-à-dire trouveront la même valeur pour chacune des décimales : c'est là le critérium concret auquel on doit recourir, sous peine de s'égarer dans des distinctions purement théoriques, peut-être intéressantes pour les philosophes mais sans aucune valeur pratique. Il pourrait sembler inutile de démontrer ce fait, puisque l'ensemble E a été construit précisément de manière à renfermer tous les nombres α; mais il est nécessaire de répondre à l'objection de M. Richard. Ne peut-on, de l'ensemble dénombrable E précédemment défini, déduire un nombre β par la méthode qui vient d'être rappelée à propos de la suite (1); ce nombre β serait bien défini par un nombre limité de caractères, à savoir ceux qui viennent d'être écrits depuis le commencement de cette Note, et cependant il n'appartiendrait pas à l'ensemble E. Voilà l'objection. La réponse est aisée : Lorsque la définition d'un nombre β fait, comme la précédente, intervenir une infi-

([1]) *Voir* notamment : Schœnflies, *Die Entwickelung der Lehre von den Punktmannigfaltigkeiten,* zweiter Teil, § 7.

nité de nombres α précédemment définis, il faut, pour que β soit défini sans ambiguïté, que la suite des α soit elle-même définie sans aucune ambiguïté possible. Or, ce n'est manifestement pas le cas pour la définition précédente, vu qu'il est douteux si le nombre β fait ou non partie de la suite des α; les deux pages qui précèdent ne peuvent donc être regardées comme définissant sans ambiguïté un nombre β au moyen d'un nombre fini de mots : par suite, la contradiction signalée n'existe pas.

Ceci ne veut pas dire que nous excluons, de l'ensemble des définitions renfermant un nombre fini de mots, celles qui supposent une infinité dénombrable d'opérations préalables : cette exclusion serait entièrement arbitraire; mais il faut que cette infinité dénombrable puisse être formée [1] et classée sans aucune ambiguïté possible. Voici, par exemple, une définition correcte :

Formons tous les nombres algébriques réels α_n, c'est-à-dire les racines d'une équation algébrique à coefficients entiers; on sait les classer dans un ordre déterminé : pour cela, l'équation à coefficients entiers s'écrivant

$$a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n = 0,$$

on considère la somme

$$n + |a_0| + |a_1| + \ldots + |a_n|$$

qui est un nombre entier et qu'on appellera *rang de* α_n; il existe un nombre limité de α_n ayant un rang donné; on peut les ranger par ordre de grandeur croissante; il est donc aisé d'écrire les α_n sans ambiguïté sous la forme d'une suite telle que (1); le nombre α qui s'en déduit d'après la règle indiquée plus haut est donc bien déterminé; tout mathématicien peut le calculer avec une approximation indéfinie [2].

Il est inutile de donner d'autres exemples; les définitions peuvent être indéfiniment variées et compliquées : lorsqu'on parle d'un nombre fini de mots, on n'exclut pas le cas où il faudrait cent mille volumes in-folio pour définir un nombre : l'essentiel est que la définition n'entraîne aucune ambiguïté et, par suite, que chacune des séries dénombrables qui peuvent y figurer soit elle-même définie avec une entière précision. Si l'on imagine la complication que prennent forcément de telles définitions dès qu'on veut introduire un grand nombre de séries dénombrables, on se rend compte combien la prétendue *définition* qui conduit au *paradoxe* de M. Richard est incomplète et insuffisante; si l'on voulait la *réaliser*, on serait arrêté par de nombreuses difficultés : elle n'a donc aucune valeur

---

[1] Je ne répète pas que cette formation est entièrement définie par un nombre fini de mots.

[2] J'omets, pour abréger, quelques détails qu'il serait utile de préciser : on pourra convenir de ne prendre que les nombres algébriques compris entre 0 et 1 et de considérer pour chacun d'eux, parmi l'infinité d'équations qu'il vérifie, celle dont le rang est le moins élevé.

pratique. Pour la mettre en œuvre, en effet, il faudrait d'abord avoir résolu tous les problèmes mathématiques qui pourront jamais être posés : car, parmi les définitions possibles, il en est qui supposent la solution de problèmes; on peut en imaginer des exemples très variés; j'attirerai particulièrement l'attention sur les définitions telles que la suivante : le nombre π étant écrit sous la forme de fraction décimale, on remplace partout le chiffre 7 par le chiffre 8 et inversement, et l'on ne conserve le nombre ainsi obtenu que si ce n'est pas un nombre algébrique.

Cette dernière restriction, quelque vraisemblable que soit la solution négative, introduit une difficulté insurmontable dans l'état actuel de la Science; si elle est surmontée demain, on imaginera aisément des énoncés plus compliqués; on peut conclure que la formation effective de l'ensemble E des nombres qui peuvent être définis au moyen d'un nombre fini de mots n'est pas réalisable; ceci ne doit pas nous empêcher d'affirmer que cet ensemble est dénombrable au sens classique du mot, car il fait partie d'un ensemble dénombrable, à savoir l'ensemble des combinaisons qu'on peut former au moyen d'un nombre de lettres ou signes de ponctuation.

Mais l'ensemble E n'est pas *effectivement énumérable*, c'est-à-dire qu'on ne peut pas indiquer, au moyen d'un nombre fini de mots, un procédé sûr pour attribuer sans ambiguïté un rang déterminé à chacun de ses éléments. Il se présente, en effet, pour certains éléments, des difficultés particulières, dont la plupart, peut-être toutes, peuvent aisément être tournées au moyen de conventions (1); mais l'ensemble des conventions nécessaires exigerait, pour être entièrement formulé, une infinité de mots, car les difficultés se présentent évidemment une infinité de fois sous une infinité de formes différentes. Telle est la réponse qu'on doit faire au paradoxe de M. Richard et à tous les paradoxes analogues : il est impossible de discuter effectivement sur un problème dont tous les termes ne sont pas explicitement définis; car, dans les cas où la définition explicite exigerait une infinité de mots, elle est en dehors du domaine des Mathématiques.

Les considérations précédentes ne tendent à rien moins qu'à nier l'existence des ensembles qui ne sont pas dénombrables; il me semble en effet, comme je l'ai dit dans une Communication faite au IVe Congrès international des mathématiciens (Rome, avril 1908, paragraphe V de cette Note), que c'est là une notion *purement négative*. Il me paraît ressortir clairement de ce qui précède que l'ensemble des points d'une droite qui peuvent être effectivement définis d'une manière individuelle est un ensemble *dénombrable*, mais non *effectivement énumérable*.

Il n'est pas possible d'indiquer le moyen de fixer sur la droite un point unique et bien déterminé qui n'appartienne pas à cet ensemble; la propo-

---

(1) D'autres questions exigeraient, en outre, la solution de problèmes compliqués, qu'on peut cependant concevoir comme résolus dans l'avenir; je laisse de côté cette difficulté-là.

sition d'après laquelle *il y a* de tels points est vraie ou fausse suivant qu'on admet ou non la possibilité d'une infinité dénombrable de choix successifs; mais c'est là une question *métaphysique*, en ce sens que la réponse positive ou négative n'aura jamais aucune influence sur le développement de la Science : tous les points dont on pourra jamais avoir besoin dans les raisonnements auront été définis au moyen d'un nombre fini de mots; ils constituent le *continu pratique* qu'utilisent les mathématiciens.

La distinction des ensembles en ensembles dénombrables et non dénombrables me paraît donc sans valeur pratique, car tous les ensembles qu'on pourra jamais considérer sont dénombrables, *au sens classique du terme*, puisqu'ils sont des parties aliquotes d'ensembles dénombrables; mais ils ne sont pas tous *effectivement énumérables :* pour certains d'entre eux l'attribution d'un rang déterminé à chacun de leurs éléments exigerait une infinité de choix et, par suite, ne pourrait être réalisée au moyen d'un nombre limité de mots. On doit donc distinguer, au point de vue pratique, deux classes d'ensembles : les ensembles effectivement énumérables et ceux qui ne le sont pas. Une partie aliquote d'un ensemble effectivement énumérable n'est pas nécessairement effectivement énumérable : l'ensemble E considéré plus haut en fournit un exemple. C'est là la différence essentielle entre la notion d'ensemble effectivement énumérable et la notion classique dénombrable. La nouvelle notion me paraît avoir le grand avantage d'être moins métaphysique, c'est-à-dire de s'appuyer exclusivement sur des réalités observables (1).

Tous les prétendus paradoxes de la théorie des ensembles proviennent de ce qu'on a admis comme évidente la proposition suivante : *Tout ensemble dénombrable est effectivement énumérable.* Or, il résulte clairement de ce qui précède que cette proposition est inexacte.

VII. — La Philosophie mathématique et l'infini (2).

L'Ouvrage que vient de publier M. Léon Brunschvicg (3) me paraît être un des plus puissants efforts qu'ait jamais tenté un philosophe pour s'assi-

---

(1) Je ne puis entrer ici dans tous les détails qui seraient nécessaires pour faire voir que, partout où l'on croit utiliser la notion d'ensemble non dénombrable pour un but concret, le raisonnement peut être modifié de manière à n'utiliser que la notion d'ensemble *effectivement énumérable;* j'en donnerai cependant un exemple. On sait qu'on utilise fréquemment le fait qu'une série ne peut être convergente si l'ensemble de ses termes n'est pas dénombrable; on peut ajouter que cet ensemble est alors *effectivement énumérable;* car, les termes dont la valeur absolue dépasse un nombre fixe étant forcément un nombre limité, on peut indiquer au moyen d'un nombre fini de mots le moyen de les énumérer effectivement.

(2) *Revue du Mois*, août 1912.

(3) *Les Étapes de la Philosophie mathématique* (F. Alcan, 1912).

miler une discipline aussi étendue que la science mathématique et pour essayer de traduire en langage philosophique les résultats acquis par les savants. Ce Livre ne peut être analysé en quelques pages; tous ceux qui s'intéressent à la philosophie des sciences ne manqueront pas de le lire et de le méditer.

Je voudrais discuter ici un point particulier sur lequel il me semble que M. Brunschvicg a adopté trop aisément, au sujet de mon attitude sur la question de l'infini et du transfini, une manière de voir de M. Winter; l'importance même de l'Ouvrage de M. Brunschvicg me paraît rendre nécessaire de signaler ce qui me semble un malentendu. Peut-être, d'ailleurs, la discussion de ce point particulier jettera-t-elle quelque lumière sur les différences de point de vue, peut-être inévitables entre philosophes et hommes de science.

Qu'on me permette de citer tout d'abord M. Brunschvicg (1) :

« ..... S'il fallait voir dans l'attitude empiriste autre chose qu'une règle de prudence particulièrement opportune et frappante de la part de ceux-là mêmes qui ont manié les notions cantoriennes et les ont fait fructifier sur le champ de la Mathématique positive, si l'on devait l'ériger en une métaphysique qui s'enfermerait rigoureusement dans l'enceinte du fini, on soulèverait des difficultés pratiques qu'on ne parviendrait guère à écarter que par des expédients de langage. Par exemple M. Borel borne le domaine de la science positive aux ensembles effectivement énumérables, c'est-à-dire tels qu'on puisse « indiquer au moyen d'un nombre fini de mots un procédé sûr pour attribuer sans ambiguïté un rang déterminé à chacun de leurs éléments (2) » et M. Poincaré donne cette règle de ne « jamais envisager que des objets susceptibles d'être définis en un nombre fini de mots (3) ». Seulement cette règle si simple prête dans l'application à plus d'une question litigieuse : « Car, parmi les mots en nombre fini, objecte avec toute raison M. Winter, pourront se rencontrer des expressions telles que *aussi souvent qu'on veut*, *aussi grand qu'on veut*, ou des mots tels que *toujours*, *indéfiniment*, *infiniment;* sommes-nous bien sûrs, dans ce cas, que la définition n'implique pas, au moins en puissance, l'infinité de mots qu'il faudrait éviter (4) ? ».

La forme donnée par M. Winter à son objection me paraît caractéristique de la tournure d'esprit philosophique, s'opposant à la tournure d'esprit scientifique. Je ne crois pas en effet qu'aucun mathématicien se demande jamais, *pendant qu'il fait des mathématiques*, si une définition en termes finis implique ou non une infinité de mots. Il n'y a pour lui qu'une distinction utile : certaines définitions permettent de connaître

(1) *Loc. cit.*, p. 532.

(2) *Les « paradoxes » de la théorie des ensembles* (*Annales scientifiques de l'École Normale*, 1908, p. 446).

(3) *La logique de l'infini* (*Revue de Métaphysique*, 1909, p. 482).

(4) *Note sur l'infini en Mathématiques* (*ibid.*, 1911, p. 615).

parfaitement l'être mathématique dont on parle, de calculer sur lui sans aucune ambiguïté ou confusion possible, d'autres ne le permettent pas.

Si l'on me demande de donner des exemples de ces dernières définitions, j'avoue que je serai fort embarassé, car je n'ai jamais pu bien comprendre l'attitude de ceux qui les admettent, et je crains d'exposer si mal cette attitude que la justification de l'attitude opposée, qui est la mienne, n'apparaisse comme véritablement trop aisée. Je vais cependant essayer de mon mieux de ne pas trahir la pensée de géomètres tels que M. Zermelo et M. Hadamard (1), en prenant un exemple aussi simple que possible. Considérons un nombre décimal illimité, par exemple celui qu'on désigne habituellement par la lettre π;

$$\pi = 3,1415926535\ldots.$$

Ce nombre particulier est entièrement *défini* si l'on dit : « π est le rapport de la circonférence au diamètre »; dans tous les Traités de Géométrie élémentaire se trouvent indiquées des méthodes qui permettent, au moyen de cette définition, de calculer autant de décimales qu'on veut. Ce calcul serait naturellement fort long, si l'on exigeait un très grand nombre de décimales; mais nous concevons qu'il puisse être effectué; nous pouvons d'ailleurs espérer que le perfectionnement de l'analyse permette d'abréger la durée des calculs; mais, quoi qu'il en soit de ce dernier point, il paraît incontestable que la valeur de la 1000e ou de la 3645e décimale de π est actuellement bien déterminée, même si le calcul n'a pas encore été effectué. Beaucoup de nombres décimaux illimités sont dans le même cas (2). Sur ces points, je crois qu'on sera d'accord. On sera d'accord aussi sur le point suivant : Il est possible de définir un nombre décimal limité en demandant à un millier de personnes d'écrire chacune un chiffre au hasard; on aura ainsi un nombre bien déterminé, si ces personnes étant rangées sur une file, chacune écrit à son tour un chiffre nouveau à la suite des chiffres déjà écrits par les personnes qui la précèdent. Voici maintenant où le désaccord commence : Est-il possible de *définir* un nombre décimal illimité par un procédé analogue à celui-là ? Je ne suppose pas qu'on ait jamais songé à demander à une *infinité* de personnes d'écrire chacune un chiffre au hasard; mais M. Zermelo et M. Ha-

(1) On trouvera, dans un article de Jules Tannery cité par M. Brunschvicg (p. 530), des idées qui ne sont pas très différentes de celles que je combats; mais cet article remonte à 1897 et les idées de Jules Tannery s'étaient modifiées sur ce point bien des années avant sa mort : c'est une des raisons pour lesquelles on n'a pas cru devoir insérer cet article dans le recueil *Science et Philosophie*, car il ne l'y eût pas inséré lui-même sans des notes complémentaires dont il m'avait parlé souvent, mais que lui seul aurait pu écrire.

(2) Les lecteurs désireux de renseignements plus techniques pourront se reporter à mon Mémoire qui vient de paraître : *Le calcul des intégrales définies* (*Journal de M. Jordan*, 1912) (Note VI de ce Livre). *Voir* aussi mes *Leçons sur la théorie de la croissance* (Chap. V).

damard pensent, je crois, qu'il est possible de regarder un tel choix comme réalisé d'une manière parfaitement déterminée et que, par suite, il est possible de parler d'*un* nombre décimal ainsi défini; bien que la définition complète d'un tel nombre exige évidemment une infinité de mots. Pour ma part, je considère qu'il est possible de se poser des *problèmes de probabilités* sur *les* nombres décimaux qui seraient ainsi obtenus en choisissant les chiffres, soit entièrement au hasard, soit en imposant certaines restrictions, restrictions laissant subsister en ce choix une part de hasard, mais qu'il est impossible de parler d'*un* de ces nombres, pour la raison que, si on le désigne par a, deux mathématiciens différents, lorsqu'ils parleront de a, ne seront jamais sûrs de parler du *même* nombre.

Ce point de vue est le seul qui me paraisse mériter le nom de *réaliste;* ce n'est pas le réalisme du philosophe qui se préoccupe de savoir s'il n'y a pas de pièges dans les termes de son discours, si l'infini dont il a peur (ou dont il a soif) comme d'un dieu inconnu ne s'y cache point quelque part, c'est le réalisme du mathématicien qui regarde comme réels les êtres avec lesquels il vit quotidiennement, et dont il parle couramment, sans avoir été jamais exposé à un malentendu. Par exemple, si je considère une série *dont chaque terme est la moitié du précédent :*

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots,$$

toute personne ayant quelque culture mathématique comprend ce que signifient les points qui suivent le signe $+$; et nul ne doute que la somme de cette série ne soit 2. Un nombre fini de mots a suffi pour que nous nous entendions parfaitement; il en serait de même si nous parlions de la valeur de telle intégrale définie ou de la solution de telle équation différentielle. Le fait que le mot *infini* est employé dans une définition comme celle de l'intégrale n'enlève rien à la réalité des calculs et des raisonnements, dans tous les cas du moins où cet *infini* est énumérable, c'est-à-dire se réduit à la considération de la suite illimitée des nombres entiers, suite dont tous les mathématiciens ont, *pratiquement*, une idée suffisamment nette et claire pour qu'ils soient certains de s'entendre entre eux lorsqu'ils en parlent.

J'ai posé, il y a déjà longtemps [1], la question suivante : Est-il possible de donner au mot *transfiniment* un sens aussi universel qu'au mot *indéfiniment*, de telle manière qu'on puisse parler avec la même sécurité d'une suite transfinie qu'on le fait d'une suite indéfinie ? J'ai d'ailleurs posé cette question d'une manière qui n'était nullement tendancieuse, hésitant fort à cette époque entre la positive et la négative. Aujourd'hui j'incline vers la négative, mais il importe de remarquer que la réponse positive ne modifierait pas mon attitude dans la question essentielle pour

---

[1] *A propos de « l'infini nouveau »; l'antinomie du transfini* (*Revue philosophique*, 1899).

moi du *nombre fini de mots*. Peut-être n'est-il pas inutile d'insister un peu sur ce point en rappelant les définitions des nombres tranfinis.

Lorsqu'on écrit une suite telle que la suite naturelle des nombres entiers

$$1,\quad 2,\quad 3,\quad 4,\quad \ldots,$$

on se la représente sous la forme d'une série illimitée simple; il peut être commode d'avoir un nom abrégé pour toute série illimitée qui se présente sous cette forme simple; M. Georg Cantor a proposé le symbole ω qu'il appelle le premier nombre transfini. Si l'on suppose que les trois premiers nombres 1, 2, 3 sont reportés sur une seconde ligne, par suite regardés comme *après* tous les nombres de la première ligne, nous aurons la disposition suivante :

$$\begin{array}{llllll} 4, & 5, & 6, & 7, & 8, & \ldots, \\ 1, & 2, & 3, & & & \end{array}$$

qu'on désignera par $\omega + 3$. Si l'on écrit d'abord tous les nombres impairs, puis tous les nombres pairs, ce qui donne

$$\begin{array}{llll} 1, & 3, & 5, & \ldots, \\ 2, & 4, & 6, & \ldots, \end{array}$$

on aura $\omega + \omega$, c'est-à-dire 2ω.

On peut aussi avoir une infinité de lignes, en écrivant par exemple sur la première les nombres impairs, puis sur la seconde les doubles de ces nombres, sur la troisième les doubles des nombres de la seconde, etc. :

$$\begin{array}{lllll} 1, & 3, & 5, & 7, & \ldots, \\ 2, & 6, & 10, & 14, & \ldots, \\ 4, & 12, & 20, & 28, & \ldots, \\ \ldots & \ldots & \ldots & \ldots & \ldots \end{array}$$

Le Tableau ainsi formé comprend ω lignes dont chacune est notée ω; on le notera ω^2.

Sans qu'il soit besoin de nouvelles explications, on voit qu'un Tableau tel que le suivant

$$\begin{array}{lllll} 2, & 6, & 10, & 14, & \ldots, \\ 4, & 12, & 20, & 28, & \ldots, \\ \ldots & \ldots & \ldots & \ldots & \ldots, \\ 5, & 7, & 9, & \ldots, & \\ 1, & 3 & & & \end{array}$$

sera noté $\omega^2 + \omega + 2$.

On voit qu'un nombre transfini n'est pas autre chose qu'un procédé de rangement des nombres entiers, procédé qui est parfaitement défini au moyen d'un nombre fini de mots. La difficulté est la suivante : on peut, au moyen d'un nombre fini de mots, définir beaucoup de nombres transfinis,

et indiquer les notations par lesquelles on les désignera. Mais on ne peut pas les définir *tous*. Ceci veut dire que, quelles que soient les définitions qu'on aura données (définitions comprenant un nombre fini de mots parmi lesquels peut figurer, bien entendu, le mot indéfiniment, dont le sens mathématique est clair, puisqu'il dérive de la suite naturelle des nombres entiers), il y aura des nombres transfinis qui échapperont à ces définitions, mais qui pourront être définis par d'autres mots en nombre fini, sans qu'on arrive jamais au bout.

On ne saurait objecter à ceci qu'il en est de même pour la suite des nombres entiers; la notation décimale, qui est expliquée dans tout Traité d'Arithmétique, donne en effet pour *tout* nombre entier une écriture unique, bien déterminée, sur laquelle aucune ambiguïté n'est possible : il a donc suffi d'un nombre fini de mots pour que tous les mathématiciens soient assurés qu'ils écriront de la même manière deux nombres entiers égaux. Il n'y a rien de pareil pour les nombres transfinis; il faudrait une infinité de mots pour en fixer la notation d'une manière dépourvue d'ambiguïté. Je crains bien que cette difficulté ne soit insoluble : elle est en tout cas insoluble par voie analytique; elle ne pourrait être résolue que si le mot *transfiniment*, par une synthèse hardie, acquérait un sens aussi clair et aussi universellement dépourvu d'ambiguïté que le sens actuellement adopté pour le mot *indéfiniment*. Je ne reviendrai pas ici sur les raisons pour lesquelles, si l'on admet que le transfini renferme quelque chose de plus que le théorème de Paul du Bois-Reymond, on doit en conclure que le mot *indéfiniment* renferme quelque chose de plus que la remarque finitiste : après chaque entier, il y en a un autre (¹).

Je pense que la position « empiriste » ou « réaliste » que j'ai adoptée paraîtra maintenant dépourvue de toute contradiction. Il m'importe peu que les mots employés dans une définition impliquent ou non une infinité d'opérations; la seule question est de savoir si la loi de ces opérations peut être donnée au moyen d'un nombre fini de mots, de telle manière que cette loi soit conçue de la même manière par tous les mathématiciens, et que le résultat soit par suite défini d'une manière unique sans ambiguïté possible. S'il en est ainsi, on fait de la science; s'il n'en est pas ainsi, on est, à mon avis, hors du domaine de la science.

Bien entendu, il est possible de raisonner sur une classe d'êtres mathématiques, par exemple sur tous les nombres réels, ou sur toutes les fonctions continues, cette classe étant définie au moyen d'un nombre fini de mots, bien que les individus ne puissent être *tous* définis ainsi. On obtient ainsi les propriétés générales de la *classe*, mais on ne raisonne jamais en réalité sur un individu *déterminé* de cette classe, à moins qu'il ne s'agisse d'un individu très particulier qui a pu être distingué de tous les autres par une définition finie.

On pourrait se demander si tout cela n'est pas une querelle purement

(¹) *Voir* les articles cités de la *Revue philosophique*.

verbale, car enfin personne n'a jamais écrit une infinité de mots : les idéalistes les plus impénitents n'ont jamais *réellement* considéré une définition non finie; ils en sont réduits à faire la même chose que les réalistes, même s'ils pensent qu'ils font autre chose : lorsqu'ils disent qu'ils raisonnent sur un nombre incommensurable arbitraire, mais déterminé, s'ils ne le déterminent pas par un nombre fini de mots, ils ne peuvent que faire des raisonnements qui s'appliquent aussi à un nombre indéterminé, c'est-à-dire à tous les nombres d'une certaine classe; le réaliste peut donc interpréter leur raisonnement comme s'appliquant, non au nombre dont il conteste la définition, mais à la classe tout entière.

J'aurais volontiers adopté ce point de vue sceptique, et renoncé à écrire ces pages, regardant l'attitude réaliste comme n'ayant pas besoin d'être défendue, s'il ne m'avait semblé que ces discussions sur l'infini sont de nature à jeter quelque lumière sur des problèmes dont tous les éléments sont finis. Il peut paraître paradoxal d'expliquer le fini par l'infini, considéré comme plus simple; c'est cependant un procédé fort courant en Physique mathématique : on remplace une suite discontinue d'éléments par un solide ou une surface continus; car il est plus simple d'étudier le continu divisible à l'infini que le discontinu divisible en un nombre extrêmement grand de parties. Actuellement, dans les théories de physique moléculaire, on fait volontiers usage simultanément de variables continues et de variables discontinues. Il est donc naturel de rechercher ce que deviennent les considérations précédentes si au nombre infini on substitue un nombre *fini très grand*.

Un gramme de matière renferme un nombre de molécules qui se chiffre par millions de milliards de milliards (c'est-à-dire est de l'ordre de 10^{24}); pour étudier d'une manière parfaite ses propriétés, il faudrait tout d'abord connaître d'une manière précise les positions et les vitesses de toutes les molécules, ce qui exigerait qu'on se donne des millions de milliards de milliards de nombres. Il est pratiquement aussi impossible d'écrire tous ces nombres qu'il est théoriquement impossible d'en écrire une infinité. La nécessité théorique des définitions au moyen d'un nombre fini de mots se transforme donc, sur le terrain des applications physiques, en la nécessité pratique des définitions au moyen d'un petit nombre de mots : ces définitions seules peuvent être considérées comme étant scientifiques. Et si la nature n'est pas ainsi? Alors la science est impossible. Mais tout porte à croire que la science est possible, et que les propriétés physiques qui nous intéressent peuvent, même si elles dépendent en apparence d'un nombre colossal de variables, être suffisamment définies au moyen d'un nombre relativement faible de données. Ces données seront généralement des valeurs moyennes, accompagnées d'indications sur les écarts statistiques par rapport à ces valeurs moyennes; on pourrait d'ailleurs concevoir, d'un point de vue abstrait, que ces écarts suivent une loi entièrement différente de la loi de Gauss; mais cette loi devrait pouvoir être explicitement formulée (par un petit nombre de mots) sous peine d'être pratiquement comme si elle n'existait pas.

Il n'est de science possible que des phénomènes qui peuvent être décrits d'une manière suffisante au moyen d'un nombre de mots accessible à un homme pendant la durée de sa vie; c'est là peut-être la raison profonde pour laquelle certains phénomènes vitaux et psychologiques échappent dans leur détail à l'investigation scientifique, la raison pour laquelle il sera sans doute toujours impossible de déduire de l'autopsie du cerveau d'un homme la connaissance des pensées qu'aurait eues cet homme s'il n'était point mort. Mais sans s'égarer aussi loin de la science actuelle, il est permis de se demander s'il peut exister des phénomènes physiques dont la connaissance complète n'exige qu'un nombre fini relativement faible de données, tandis que leur explication exigerait plus de mots que l'intelligence d'un homme ne pourrait en saisir. De tels phénomènes échapperaient à la science humaine, tout au moins à la science « toute subjective » qui me paraît être la seule science dont on puisse légitimement parler. J'avoue en effet ne pas comprendre ce que veut dire M. Hadamard lorsqu'il écrit [1] : « Je dirais plutôt : l'ordination est-elle possible ? (et non pas même peut-*on* ordonner, de crainte d'avoir à penser à ce qu'est *on*); Baire dirait : pouvons-*nous* ordonner ? Question toute subjective, à mon avis. » Je ne conçois pas en effet ce que peut être la possibilité en soi d'un acte qui serait impossible pour tout esprit humain; c'est pour moi une pure abstraction métaphysique, en dehors de toute réalité scientifique.

On se trouve ainsi conduit à se poser un problème qui est peut-être le plus important de ceux que soulèvent les théories physiques modernes : certaines sciences ne sont-elles pas sur le point d'atteindre les phénomènes dont l'explication scientifique détaillée dépasse les possibilités humaines, simplement parce qu'elle exigerait trop de mots. Je ne veux point aborder ici cette question [2] : je rappellerai simplement les études pénétrantes de M. Jean Perrin qui ont paru ici même [3]; le mouvement brownien n'est-il pas précisément l'un de ces phénomènes qui peut être décrit, sous sa forme générale, au moyen d'un nombre relativement faible de mots, mais dont la prévision détaillée exigerait un nombre physiquement « infini » d'équations, autant d'équations qu'il y a de molécules; on doit se contenter d'une prévision statistique.

Nous sommes en apparence bien loin de notre point de départ; je pense cependant que la relation entre l'infini mathématique et ce qu'on peut appeler l'infini pratique des physiciens modernes (nombre très grand des molécules) n'est pas une conception purement artificielle; on sera amené, par le développement des théories physiques, à approfondir cette relation,

---

(1) *Voir* plus haut, p. 157.

(2) *Voir* mon article sur *Le hasard et la vérité scientifique* (*Revue de Paris*, 1er août 1912).

(3) *La discontinuité de la matière* (*Revue du Mois*, 10 mars 1906, t. I, p. 323); *Peut-on peser un atome avec précision ?* (*Revue du Mois*, 10 novembre 1908, t. VI, p. 513).

et c'est peut-être dans l'étude de ces théories que les spéculations analytiques sur l'infini et le discontinu trouveront leur signification complète et leur champ d'application le plus large.

VIII. — L'Infini mathématique et la réalité (1).

A propos de l'article précédent, j'ai reçu une lettre de mon ami Jacques Hadamard, à la suite de conversations dans lesquelles nos points de vue s'étaient opposés. Voici cette lettre :

« Cher ami,

» Je reviens encore une fois sur la question de l'empirisme et de l'idéalisme, à propos de ton article de la *Revue du Mois* (août 1912).

» Ce n'est pas que j'aie à m'élever plus spécialement contre celui-là que contre les précédents. Bien au contraire il ne me semble pas avoir à me plaindre des conclusions qui s'en dégagent. Au fond, elles réduisent le différend à bien peu de chose.

» D'abord, il en ressort que, si le raisonnement de Zermelo n' « existe » pas, le mouvement brownien n'existe pas non plus, ou guère. Il me semble que, pour le « choix » de Zermelo, une existence du même ordre que celle du mouvement brownien est encore assez acceptable.

» Mais nous sommes encore bien plus près d'être d'accord lorsque tu précises en ces termes : « Il est possible de raisonner sur une classe d'êtres » mathématiques, par exemple sur tous les nombres réels, ou sur toutes » les fonctions continues, cette classe étant définie au moyen d'un nombre » fini de mots, bien que les individus ne puissent être *tous* définis » ainsi. »

» Je ne suis pas bien sûr d'avoir jamais dit autre chose. Car, enfin, il faut que tous ces individus existent en quelque manière, pour former la classe.

» En tout cas, à ce compte, si un mouvement brownien n'existe pas, on a le droit de parler de la classe des mouvements browniens. (Je ne sais si ce serait le vrai moyen d'y voir clair : je ne sais si, une fois rédigées ainsi, les *Leçons sur la théorie des gaz* resteraient ainsi utiles.) De même, par conséquent, Zermelo aurait démontré, sinon qu'il existe *une* manière de bien ordonner le continu, du moins qu'il existe une *classe* de pareilles ordinations. C'est toujours cela : va pour la classe !

» Cela veut dire, en somme, que (si le raisonnement de Zermelo n'est pas ultérieurement perfectionné) nous ne pourrons jamais raisonner que sur les propriétés communes à *toutes* ces ordinations. Je le crois volontiers. Il y a bien d'autres choses que nous ne pourrons jamais connaître.

» C'est même pour cela que, comme tu le dis fort bien, la Science est

---

(1) *Revue du Mois*, 10 juillet 1914.

subjective. Elle est résignée à cela depuis longtemps. Seulement, il est des savants, dont je suis, qui croient que, par contre, son objet est objectif, si tu me permets ce truisme apparent. Celui-ci est d'ailleurs si loin d'être réel qu'il résume, au contraire, ce qui reste du débat entre les empiristes et les.... — je ne peux dire : les idéalistes, car, comme l'a fait remarquer Poincaré dans son article de *Scientia*, le point de vue habituellement désigné sous ce nom par les philosophes n'est autre, ici, que le point de vue empiriste. Ce n'est même pas là le côté le moins curieux du débat : c'est la première fois que l'idéalisme classique, habitué à s'attaquer au monde physique et sensible, se trouve transporté dans le domaine mathématique.

» Enfin, n'oublions pas qu'il y a un autre point sur lequel nous tombons d'accord, si je ne me trompe : c'est la nécessité d'adopter le point de vue auquel je me place, toutes les fois qu'il s'agit d'intuition et de découverte (la Science mathématique n'a jamais progressé autrement, depuis la création du calcul infinitésimal). Et cela me donne l'espérance d'avoir l'avenir pour moi, puisque la rigueur n'a jamais eu pour objet que de sanctionner et de légitimer les conquêtes de l'intuition.

» A toi.

» J. Hadamard. »

Le point sur lequel je m'accorderais le plus volontiers avec M. Hadamard, c'est sur la possibilité même d'un accord ; je ne puis en effet, malgré l'opinion de Poincaré, croire que cet accord puisse être impossible entre mathématiciens, en raison de la différence des natures de leurs esprits; du moment qu'il s'agit de mathématiques et non de philosophie, le désaccord ne peut provenir que d'un malentendu. Je voudrais essayer d'éclaircir ce malentendu en me plaçant à un point de vue aussi objectif que possible, c'est-à-dire en essayant de distinguer le symbolisme mathématique de l'interprétation qu'en donnent les divers auteurs et lecteurs.

On a souvent cité, en ces dernières années, un passage de M. Hilbert qui, à propos des principes de la Géométrie, indique nettement le point de vue axiomatique ou symbolique : « Pensons trois systèmes de *choses*, que nous appellerons points, droites et plans... ». On attribue *a priori* à ces *choses* certaines propriétés, assujetties simplement à ne pas être contradictoires en soi, et la déduction de toutes les propriétés qui en résultent constitue une géométrie. Cette manière de procéder est évidemment légitime, en ce sens que l'on a toujours le droit de créer un vocabulaire et de construire avec ce vocabulaire un édifice logique ; mais l'absence de contradiction logique ne suffit pas à caractériser la construction scientifique. Si les travaux géométriques de M. Hilbert sont considérés comme faisant partie de la science mathématique, c'est en raison des relations étroites entre les « choses » que M. Hilbert appelle points, droites et plans et les « choses » que le vulgaire appelle points, droites et plans. Je n'ai pas à insister ici sur les analogies ni sur les différences entre ce que l'on peut appeler la

géométrie logique ou arithmétique et la *géométrie expérimentale ;* c'est là une question qui peut prêter à d'intéressantes discussions philosophiques ; ce qui paraît hors de toute discussion, c'est que M. Hilbert n'aurait jamais songer à doter certaines « *choses* » des propriétés axiomatiques qu'il a codifiées si la géométrie n'eût pas existé depuis longtemps sous sa forme usuelle. Au sujet de la géométrie classique et de ses rapports avec la réalité, je ferai encore une remarque évidente : personne ne croit qu'il existe des objets satisfaisant rigoureusement aux définitions géométriques, mais la géométrie n'aurait pas été créée si les hommes ne croyaient pas tous, au moins dans la pratique, à l'existence d'objets qui réalisent approximativement ces définitions.

De même qu'il est loisible à M. Hilbert de penser à des « choses » qu'il appelle *points droites*, etc., de doter ces choses des propriétés qu'il lui plaît (non contradictoires) et de raisonner sur ces propriétés, il ne saurait être interdit à M. Hadamard [ou à M. Cantor (1), ou à M. König (2), ou à M Schœnflies (3), ou à M. Zermelo, ou à M. Hausdorff (4)] de penser à des « choses » qu'on appellera infini, ou puissances successives, ou aleph, ou aleph exposant aleph, etc., et de doter ces « choses » de n'importe quelles propriétés, du moment que ces propriétés ne sont pas contradictoires. Le système logique ainsi constitué est irréprochable au point de vue logique, du moment que les règles de la logique ont été observées dans la déduction, et il n'est pas douteux que si les mots « continu » et « ensemble bien ordonné » ont des définitions convenables, la proposition « le continu est un ensemble bien ordonné » ne pourra pas être mise en doute.

Je suis tout à fait certain que les mathématiciens qui ont élaboré ces constructions logiques, et qui avaient fait preuve en d'autres domaines des qualités nécessaires de rigueur, ont conservé ces qualités logiques et, s'ils se sont trompés (cela peut arriver à tout le monde), ne se sont trompés que fort rarement. Sans qu'il soit nécessaire d'examiner minutieusement en détail leurs déductions et leurs calculs, j'attribue donc à leurs conclusions la même valeur que nous sommes habitués à attribuer aux conclusions d'un Mémoire mathématique dont l'auteur nous est connu comme ne se trompant qu'exceptionnellement.

Mais il faut bien comprendre qu'à ce point de vue logique la valeur des conclusions est purement verbale, et que l'édifice construit est pour le moment sans aucun rapport avec la réalité, même si l'on accepte de dési-

(1) G. Cantor, *Grundlagen einer allgemeinen Mannigfaltigkeitslehre* (Leipzig, 1883).

(2) J. König, *Neue Grundlagen der Logik, Arithmetik und Mengenlehre* (Leipzig, Veit, 1914).

(3) Schœnflies, *Entwicklelung der Mengenlehre und ihrer Anwendungen* (Teubner, 1913).

(4) Hausdorff, *Grundlagen der Mengenlehre* (Leipzig, Veit, 1914).

gner sous le nom de *réalité* les parties classiques des mathématiques. Nous n'accordons donc rien en accordant la valeur logique des déductions où l'infini, les alephs, etc., sont définis d'une manière axiomatique; il reste à examiner si ces définitions axiomatiques sont en accord avec les idées ou les images qui résultent pour les mathématiciens de leur contact quotidien avec les parties classiques de leur science. Nous avons d'autant plus le droit d'exiger cet accord que tous ceux qui ont étudié l'infini ont évidemment cherché à en faire une réalité et certains d'entre eux peut-être ont considéré cette réalité comme allant de soi d'une manière si naturelle qu'ils n'ont même pas cru qu'ils procédaient par la méthode axiomatique, mais se sont figurés qu'ils donnaient des définitions d'une légitimité incontestable.

Il est cependant impossible de ne pas reconnaître que nous n'observons jamais qu'un nombre fini d'objets et ne traçons qu'un nombre fini de signes; les mathématiciens n'ont considéré et ne considéreront jamais individuellement qu'un nombre fini d'éléments d'une nature quelconque; la position des finitistes absolus, qui se refusent à admettre aucun infini, est donc logiquement très forte.

Pratiquement, elle l'est beaucoup moins. Il suffit d'avoir étudié les éléments de l'analyse mathématique pour se rendre compte à quel degré de complication conduirait la prétention d'exclure l'infini tel qu'on le considère dans ces éléments. La considération des séries infinies, pour ne citer que cet exemple, conduit à des formules à la fois bien plus élégantes et plus pratiques que ne le seraient les formules obtenues en prenant seulement un nombre fini très grand de termes. En ce cas, l'infini se présente sous la forme de l'énumérable, c'est-à-dire d'une suite analogue à la suite naturelle des nombres entiers [1]. Il paraît incontestable que les mathématiciens se font ou croient se faire de cette suite indéfinie

$$(S) \qquad 1, 2, 3, 4, 5, 6, \ldots\ldots\ldots$$

une image parfaitement claire. Ce qui est peut-être plus important encore, la considération de cette suite n'a jamais conduit à des intuitions inexactes,

---

[1] Il y aurait lieu, pour être complet, de critiquer aussi la notion du continu, que beaucoup d'excellents esprits considèrent comme donnée par l'intuition géométrique. A quel point cette intuition spatiale est défectueuse, c'est ce qui résulte avec évidence de toute étude approfondie de cette notion. Aucun esprit géométrique se basant uniquement sur l'intuition spatiale ne croira que l'on puisse laisser subsister la plus grosse partie d'une droite après que l'on a enlevé un petit segment autour de chacun des points dont l'abscisse est rationnelle; ces points étant denses sur toute portion de la droite, on ne *voit* pas comment on peut enlever un segment autour de chacun d'eux et cependant laisser quelque chose. Ce fait, et d'autres analogues, doivent rendre très prudent dans l'introduction du continu par la voie géométrique. Je me contenterai de renvoyer à des remarques que j'ai développées ailleurs; *voir* notamment *Le calcul des intégrales définies* (*Journal de Mathématiques* de M. Jordan, 1912; Note VI de cet Ouvrage).

à des paradoxes ou à des contradictions. Le seul caractère par lequel elle diffère d'une suite finie, c'est qu'elle a même puissance qu'une de ses parties aliquotes, ce qui s'éclaircit immédiatement par la considération des nombres pairs :

(S′) 2, 4, 6, 8, 10, 12, ...

Il est clair que (S′) renferme une partie des éléments de (S); il est non moins clair que l'on peut établir entre (S) et (S′) une correspondance univoque et réciproque, puisqu'il suffit pour cela de faire se correspondre les termes de même rang; chaque terme de (S′) étant le double du terme correspondant de (S).

Au sujet de la suite naturelle des nombres entiers, les finitistes peuvent cependant formuler des objections qui m'ont été exposées d'une manière particulièrement claire par M. Baire, dans une conversation qui remonte à une douzaine d'années. Il y a une certaine part d'illusion dans la clarté que l'usage de la numération décimale jette sur notre intuition de la suite naturelle des nombres entiers. Cette numération est assurément un instrument admirable au point de vue pratique, si on la compare au procédé primitif qui consisterait à figurer un nombre par autant de traits verticaux qu'il renferme d'unités. Mais il faut prendre garde que cet avantage n'est réel que pour les « petits » nombres, n'ayant pas plus de chiffres que nous ne pouvons commodément en écrire. S'il s'agissait de nombres ayant seulement quelques milliards de milliards de chiffres, l'impossibilité de les écrire serait pratiquement la même dans le système décimal que dans le système primitif des traits identiques; car il importe peu que le poids du papier nécessaire pour écrire un tel nombre soit de l'ordre du milliard de tonnes ou d'un ordre encore plus élevé.

La difficulté se recule encore davantage, mais reste au fond la même si, au lieu d'un nombre *déterminé* assez grand, on se propose simplement de figurer par des symboles simples des nombres de plus en plus grands. En introduisant des symboles convenables, il sera possible d'écrire effectivement d'une manière très simple certains nombres qui, dans le système décimal, ne pourraient pas humainement être écrits. Si l'on désigne par exemple par (n) un nombre de n chiffres, $((n))$ sera un nombre dont le nombre de chiffres a lui-même n chiffres, de sorte que $((10))$ désigne un nombre de 10 milliards de chiffres. En écrivant une parenthèse de plus, $(((10)))$ désignera un nombre dont le nombre de chiffres a lui-même 10 milliards de chiffres. On peut aller bien plus loin et imaginer que l'on écrive des milliers de parenthèses ainsi superposées, puis qu'au lieu d'écrire effectivement ces parenthèses on désigne par $[n]$ le nombre $((\ldots((10))\ldots))$ dans le cas où il y a n parenthèses; en superposant les crochets, on aura une notation encore plus condensée, et ainsi de suite. Il est impossible de ne pas être frappé de l'analogie entre la définition de tels symboles arithmétiques et la marche suivie par M. Georg Cantor pour construire les nom-

bres transfinis. Dans les deux cas, on est arrêté par la même difficulté; on ne peut formuler effectivement qu'un nombre fini de conventions et, si loin que ces conventions permettent d'aller, les nombres qu'elles permettent d'atteindre pratiquement ne sont rien à l'égard de ceux qui leur échappent. Lorsqu'on sera arrivé à fixer des conventions telles que l'on puisse décrire en quelques pages un nombre colossalement grand, il y aura des nombres encore bien plus grands qui, avec ces mêmes conventions, ne pourront être définis que par des écritures exigeant des milliards de tonnes de papier.

J'ai tenu à exposer ces objections finitistes en tâchant de leur donner toute leur force; elles ne me paraissent cependant pas décisives lorsqu'on se place, comme j'essaye de le faire, sur le terrain des réalités, c'est-à-dire lorsqu'on examine ce que font réellement ceux des mathématiciens qui ne se préoccupent que de comprendre et de faire avancer leur science, sans s'inquiéter des controverses philosophiques.

La notion de l'infini énumérable apparaît alors comme n'étant pas seulement un symbole sur lequel on calcule d'après des conventions arbitraires et précises, mais comme une possibilité limite conçue par notre imagination, de même que la droite idéale ou le cercle parfait. Nous pouvons écrire des myriades de nombres entiers de plus en plus grands et raisonner sur ces nombres; malgré les difficultés que présente la réalisation de la suite de *tous* les nombres entiers, malgré l'impossibilité matérielle de dépasser jamais un rang fini dans cette suite [1], nous en imaginons d'une manière plus ou moins claire le prolongement indéfini. Quelles sont la nature psychologique et l'origine historique de cette conception? Nous n'avons pas à résoudre ces questions difficiles; il nous suffit de constater l'accord *pratique* des mathématiciens dans l'usage qu'ils font de ces notions et l'utilité de l'imagination de l'indéfini dans les raisonnements logiques sur les symboles qui représentent l'infini. C'est en raison de cette utilité qu'il n'est pas arbitraire d'employer les mots qui se rattachent à la notion de l'infini énumérable, de même qu'il n'est pas arbitraire d'employer les termes point, droite, plan, pour désigner les « choses » définies par M. Hilbert.

Lorsqu'il s'agit au contraire de l'infini non dénombrable, je crois constater que tous les discours par lesquels on a essayé d'en éveiller en moi l'idée n'ont jamais suggéré à mon imagination autre chose que l'infini énumérable; les raisonnements sur les symboles alephs conservent donc pour moi un caractère purement abstrait, ne correspondant à aucune réalité.

---

(1) Il est clair, en effet, que si perfectionnés que l'on puisse concevoir les systèmes de notations analogues à ceux de la page 178 qui ont été, sont ou seront imaginés par les hommes, les nombres ainsi définis effectivement, c'est-à-dire dont la définition sera écrite ou pensée par un homme, seront en nombre fini, si l'on assigne une durée finie à l'humanité.

Je ne prétends pas, bien entendu, que d'autres ne puissent avoir, maintenant ou dans l'avenir, une imagination moins pauvre que la mienne; ce que je crois pouvoir affirmer, c'est que cette imagination de l'infini non énumérable n'est pas commune chez les mathématiciens et ne joue par suite actuellement aucun rôle en Mathématiques. Mènera-t-elle un jour à des découvertes ceux qui la possèdent ou croient la posséder? Ce jour-là, je serai d'accord avec M. Hadamard; mais, pour l'instant, j'ai avec moi les mathématiciens auxquels on doit les progrès réels effectués en Mathématiques grâce à la théorie des ensembles, et notamment M. Lebesgue, dont les travaux ont eu des répercussions si importantes sur toutes les parties de l'analyse classique (¹).

Il y a une réelle analogie entre les controverses sur l'infini mathématique et les querelles entre les écoles de physiciens qu'on appelait il y a quelques années les atomistes et les énergétistes. Il ne faudrait pas croire que les récents travaux expérimentaux mettant en évidence la « réalité » des atomes (²) ont clos cette querelle. L'impossibilité humaine d'écrire les équations différentielles des mouvements de *tous* les atomes ne dépend en effet d'aucune constatation expérimentale; elle est la même aujourd'hui qu'il y a vingt ans. Si l'on admet que le but de la science est de prévoir les phénomènes observables au moyen des phénomènes observables, dans tous les cas où cette prévision est possible (³), il n'est pas douteux que ces divers phénomènes et leurs relations mutuelles devront pouvoir s'exprimer au moyen d'un nombre, non seulement fini, mais relativement petit, de paramètres et d'équations. Ces équations et les raisonnements qui les accompagnent sont en définitive des relations entre des mots et des symboles dont il est possible de donner une définition abstraite, tout comme de l' « infini » de M. Cantor ou des « droites » de M. Hilbert; au point de vue logique, l'exactitude des raisonnements et des équations ne dépend pas du sens attaché aux mots et aux symboles, ce

---

(¹) M. Lebesgue m'a communiqué, pendant que je corrigeais les épreuves de cet article, les épreuves de sa discussion avec M. Hadamard au Congrès de philosophie mathématique d'avril 1914, qui doit paraître dans la *Revue de Métaphysique et de Morale*. Il se place à un point de vue assez différent de celui que j'ai adopté ici, mais je m'accorderais avec lui sur beaucoup plus de points qu'avec M. Hadamard.

(²) Émission des particules α observées individuellement, expériences de M. Jean Perrin sur le mouvement brownien, théorie du bleu du ciel, etc. *Voir* le Livre de M. Jean Perrin, *Les Atomes*, où les preuves de la « réalité » moléculaires ont été exposées d'une manière aussi complète et aussi claire que possible.

(³) On sait que, dans certains cas, les phénomènes ne peuvent pas être prévus individuellement d'une manière précise et que l'on doit se contenter d'une prévision statistique; *voir* mon *Introduction géométrique à quelques théories physiques* (Gauthier-Villars), notamment Note VII, et mon Livre sur *Le Hasard* (F. Alcan).

sens peut du moins être varié de bien des manières : il suffit que les définitions logiques de ces mots, de ces symboles et de leurs relations mutuelles soient exemptes de contradiction. On peut ainsi concevoir une théorie logiquement parfaite et sans rapport avec la réalité; cette théorie pourrait être également interprétée dans un langage atomiste ou dans un langage où il ne serait pas question d'atomes.

Si, cependant, on constate que les raisonnements sont singulièrement facilités par l'imagination des atomes, on sera conduit à attribuer à ces atomes une « réalité ». On peut d'ailleurs se demander dans quelle mesure cette imagination est illusoire, et en particulier douter que nous arrivions à imaginer effectivement et non pas seulement verbalement les chocs innombrables des molécules. A un autre point de vue, on peut observer que la mécanique statistique n'étudie jamais un modèle mécanique déterminé, mais seulement une classe de tels modèles; doit-on attribuer une réalité au modèle unique ou seulement à la classe?

Ce sont là des questions qui me paraissent fort intéressantes et sur lesquelles j'espère revenir dans une autre occasion; je tenais seulement à signaler en passant les analogies qui les rattachent aux controverses sur l'infini mathématique.

Pour conclure, je suis parfaitement d'accord avec M. Hadamard pour penser que les intuitions qui conduisent à des découvertes sont l'essentiel en Mathématiques; mais c'est précisément pour cela que je me suis toujours efforcé de séparer celles des parties de la théorie des ensembles qui ont effectivement contribué au progrès de la théorie des fonctions, des constructions logiques purement verbales dans lesquelles on jongle avec des symboles auxquels ne correspond aucune intuition.

NOTE V.

LES PROBABILITÉS DÉNOMBRABLES ET LEURS APPLICATIONS ARITHMÉTIQUES.

I. — Les probabilités dénombrables.

1. On distingue généralement, dans les problèmes de probabilités, deux catégories principales, suivant que le nombre des cas possibles est fini ou infini : la première catégorie constitue ce qu'on appelle les *probabilités discontinues*, ou probabilités dans le domaine du discontinu, tandis que la seconde catégorie comprend les *probabilités continues* ou *probabilités géométriques*. Une telle classification apparaît comme incomplète, lorsqu'on se reporte aux résultats acquis dans la théorie des ensembles; entre la puissance des ensembles finis et la puissance du continu se place la puissance des ensembles dénombrables; je me propose de montrer brièvement l'intérêt qui s'attache aux questions de probabilités dans l'énoncé desquelles interviennent de tels ensembles; je les appellerai, pour abréger, *probabilités dénombrables*.

2. Avant de définir plus précisément les probabilités dénombrables, je voudrais indiquer en quelques mots les raisons pour lesquelles il m'a semblé que leur étude ne devait pas être plus longtemps négligée. La principale de ces raisons est l'importance de la notion d'ensemble dénombrable; cette importance n'est contestée par aucun mathématicien; mais elle me paraît être plus grande encore qu'on ne le croit généralement.

Beaucoup d'analystes, en effet, mettent au premier rang la notion du continu; c'est elle qui intervient d'une manière plus ou moins explicite dans leurs raisonnements. J'ai indiqué récemment [1] en quoi cette notion du continu, considéré comme ayant une puissance supérieure à celle du dénombrable, me paraît être une notion purement négative, la puissance des ensembles dénombrables étant la seule qui nous soit connue d'une manière positive, la seule qui intervienne *effectivement* dans nos raisonnements. Il est clair, en effet, que l'ensemble des éléments analytiques susceptibles d'être réellement définis et considérés ne peut être qu'un ensemble dénombrable; je crois que ce point de vue s'imposera chaque jour

[1] Borel, *Les paradoxes de la théorie des ensembles* (*Annales scientifiques de l'École Normale supérieure*, 3e série, t. XXV, 1908; Note IV, § VI de ce Livre).

davantage aux mathématiciens et que le continu n'aura été qu'un instrument transitoire, dont l'utilité actuelle n'est pas négligeable (comme nous allons d'ailleurs en donner des exemples dans un instant), mais qui devra être regardé seulement comme un moyen d'étudier les ensembles dénombrables, lesquels constituent la seule réalité que nous puissions atteindre.

3. Nous classerons en trois catégories les problèmes de probabilités dénombrables :

1° Le nombre des cas possibles est limité dans chaque épreuve, mais les épreuves sont en infinité dénombrable;

2° Les cas possibles sont en infinité dénombrable et le nombre des épreuves est limité;

3° Les cas possibles et les épreuves sont en infinité dénombrable.

Ces distinctions sont d'ailleurs surtout pratiques; au point de vue purement logique, il ne serait pas difficile de traiter simultanément les questions des trois catégories; cette manière de procéder me paraît présenter plus d'inconvénients que d'avantages.

4. Nous nous occuperons d'abord des problèmes de la *première catégorie*, où le nombre des cas possibles est limité dans chaque épreuve; pour éviter des écritures superflues, nous supposerons qu'il y a seulement deux cas possibles que nous appellerons, suivant un usage commode, le cas favorable et le cas défavorable; le lecteur fera facilement l'extension de nos résultats à l'hypothèse où le nombre des cas possibles est supérieur à deux, en restant fini.

On fait une infinité dénombrable d'épreuves successives, qui seront numérotées à l'aide des nombres entiers dans leur ordre naturel; la probabilité du cas favorable est désignée par p_n pour l'épreuve de rang n; la probabilité du cas défavorable est $q_n = 1 - p_n$.

Problème I. — *Quelle est la probabilité pour que le cas favorable ne se produise jamais?* Nous désignerons cette probabilité par A_0; l'application du principe des probabilités composées donne

$$A_0 = (1 - p_1)(1 - p_2)\ldots(1 - p_n)\ldots.$$

Nous discuterons tout à l'heure la légitimité de l'extension du principe au cas d'une infinité d'épreuves; étudions d'abord la formule. Si la série à termes positifs

$$(1) \qquad p_1 + p_2 + \ldots + p_n + \ldots$$

est *convergente*, il en est de même du produit infini A_0; celui-ci est alors convergent; il est différent de zéro, sauf dans le cas où l'une des probabilités p_n est égale à l'unité. *Nous exclurons ce cas :* le lecteur apercevra aisément les modifications qu'il faudrait introduire dans ce qui suit, lorsqu'on admet qu'un ou plusieurs des p_n peut être égal à l'unité. Nous

pouvons donc dire que, *dans le cas de convergence, la probabilité* A_0 *a une valeur bien déterminée, différente de zéro et de l'unité.*

Si la série (1) est divergente, il en est de même du produit A_0 : le produit des n premiers facteurs tend, lorsque n augmente indéfiniment, vers la limite *zéro;* telle est la valeur qu'on est conduit à attribuer à la probabilité cherchée A_0; quelques explications sont cependant nécessaires.

Dans le cas de la convergence, l'extension du principe des probabilités composées va de soi, et aussi la définition de la probabilité cherchée. Si l'on considère en effet les n premières épreuves, les principes classiques permettent de définir et de calculer la probabilité pour que le cas favorable ne se présente pas dans ces n épreuves; on constate que, n grandissant, non seulement cette probabilité varie très peu d'une manière *absolue*, mais varie aussi très peu d'une manière *relative*, c'est-à-dire que ses variations sont une fraction très faible de sa valeur. On peut donc, ayant assigné telle valeur qu'on veut à la précision relative qu'on désire atteindre, être certain que cette précision est effectivement atteinte au bout d'un nombre d'épreuves, peut-être fort grand, mais assignable; le passage à la limite que nous avons effectué ne soulève donc aucune difficulté et est entièrement justifié.

Il n'en est pas de même dans le cas de la divergence, et une remarque préliminaire est nécessaire. Il y a, en effet, une véritable discontinuité entre une probabilité infiniment petite, c'est-à-dire une probabilité variable qui tend vers zéro, et une probabilité égale à zéro. Si petite, en effet, que soit la probabilité du cas favorable, celui-ci est possible; tandis qu'il est impossible si la probabilité est nulle. Tels sont du moins les résultats classiques dans la théorie des probabilités discontinues; il n'en est pas de même dans la théorie des probabilités continues : la probabilité pour qu'un nombre pris *au hasard* soit rationnel est nulle [1]; cela ne veut pas dire qu'il n'y a pas de nombres rationnels. Il en sera de même pour la théorie des probabilités dénombrables : *probabilité nulle ne devra pas être considérée comme l'équivalent d'impossibilité.* Ceci étant bien entendu, il n'y a plus d'inconvénient à dire que, dans le cas de la divergence, A_0 est nul; mais on ne devra pas perdre de vue que ce langage ne signifie pas autre chose que ceci : la probabilité pour que le cas favorable ne se produise pas tend vers zéro lorsque le nombre des épreuves augmente indéfiniment.

Problème II. — *Quelle est la probabilité pour que le cas favorable se produise k fois et k fois seulement?* Nous désignerons cette probabilité par A_k. Calculons d'abord A_1. La probabilité pour que le cas favorable se produise à la première épreuve et ne se reproduise plus jamais ensuite est évidemment

$$\varpi_1 = p_1(1-p_2)(1-p_3)\dots(1-p_n)\dots.$$

---

[1] *Voir* Borel, *Remarques sur certaines questions de probabilité* (*Bulletin de la Société mathématique de France*, t. XXXIII, 1905, p. 123-128).

Dans le cas de la divergence, ϖ_1 est nul; dans le cas de la convergence, on a

$$\varpi_1 = \frac{p_1}{1-p_1} A_0.$$

De même, la probabilité ϖ_n pour que la $n^{\text{ième}}$ épreuve soit seule favorable est nulle, dans le cas de la divergence, et donnée, dans le cas de la convergence, par la formule

$$\varpi_n = \frac{p_n}{1-p_n} A_0.$$

Un raisonnement analogue à celui que nous avons fait dans le problème I permet d'appliquer, *dans le cas de la convergence*, le principe des probabilités totales; on obtient ainsi

$$A_1 = A_0 \left(\frac{p_1}{1-p_1} + \frac{p_2}{1-p_2} + \ldots + \frac{p_n}{1-p_n} + \ldots \right).$$

La série qui multiplie A_0 est visiblement convergente; nous poserons

$$u_n = \frac{p_n}{1-p_n} = \frac{p_n}{q_n}$$

et nous écrirons

$$(2) \qquad A_1 = A_0 \Sigma u_n.$$

L'extension de cette formule au cas de la divergence demande quelque précaution; la série à termes positifs Σu_n est, en effet, divergente; sa somme doit être regardée comme égale à $+\infty$ et, A_0 étant nul, A_1 se présente sous la forme $0 \times \infty$. A un autre point de vue, on peut dire que la probabilité cherchée A_1 est la somme des probabilités ϖ_n qui sont toutes nulles; mais, leur nombre n'étant pas fini, on ne peut pas en conclure sans précaution que la probabilité totale est nulle, si l'on se souvient que probabilité nulle ne signifie pas *impossibilité*. Calculons donc la probabilité σ_n pour que le cas favorable se produise une fois, et une fois seulement, sur les n premières épreuves; on a

$$\sigma_n = (1-p_1)(1-p_2)\ldots(1-p_n)(u_1+u_2+\ldots+u_n),$$

d'où l'on conclut

$$\sigma_n < e^{-(p_1+p_2+\ldots+p_n)}(u_1+u_2+\ldots+u_n),$$

et l'on aperçoit immédiatement que la divergence de la série p_n entraîne la conséquence que σ_n tend vers zéro lorsque n augmente indéfiniment; on a donc, *dans le cas de divergence*,

$$A_1 = 0.$$

On démontrerait de même que A_2 est nul, dans le cas de divergence, et

donné, dans le cas de convergence, par la formule

$$A_2 = A_0(u_1 u_2 + u_1 u_3 + u_1 u_4 + u_2 u_3 + u_1 u_5 + u_2 u_4 + \ldots + \ldots)$$

qu'on peut aussi écrire

$$A_2 = A_0 \Sigma u_n u_{n'},$$

mais en ayant soin de spécifier que la somme ne renferme que les produits de deux facteurs u_n *d'indices différents* et chacun une seule fois.

De même A_k est nul, dans le cas de divergence, et donné, dans le cas de convergence, par la formule

$$(3) \qquad A_k = A_0 \Sigma u_{n_1} u_{n_2} \ldots u_{n_k},$$

dans laquelle la signification du signe Σ ressort de ce qui précède : chaque combinaison de k indices *différents* doit y figurer *une seule fois*.

Problème III. — *Quelle est la probabilité pour que le cas favorable se produise une infinité de fois ?* Nous désignerons cette probabilité par A_∞ et allons la calculer d'abord dans le cas de convergence. Évaluons, dans ce but, la somme

$$S = A_0 + A_1 + A_2 + \ldots + A_k + \ldots.$$

Les formules (2) et (3) permettent d'écrire

$$S = A_0(1 + u_1)(1 + u_2)\ldots(1 + u_k)\ldots,$$

le produit infini étant convergent en même temps que la série (1). Mais nous avons

$$(4) \qquad A_0 = (1 - p_1)(1 - p_2)\ldots(1 - p_n)\ldots,$$

$$(5) \qquad u_n = \frac{p_n}{q_n} = \frac{p_n}{1 - p_n}$$

et, par suite,

$$1 + u_n = \frac{1}{1 - p_n},$$

$$\prod (1 + u_n) = \prod \frac{1}{1 - p_n} = \frac{1}{A_0}.$$

Nous obtenons donc finalement

$$S = A_0 \frac{1}{A_0} = 1.$$

Mais la probabilité cherchée A_∞ est évidemment égale à $1 - S$; sa valeur est donc *zéro*, résultat au sujet duquel nous ne répétons pas les remarques déjà faites.

Il est remarquable que ce résultat soit indépendant de toute hypothèse sur la *fréquence* avec laquelle doit se produire l'infinité de cas favorables.

Rien n'a été spécifié à cet égard; si l'on désigne par $\varphi(h)$ le rang de l'épreuve par laquelle se produit le cas favorable dont le rang est h, la fonction $\varphi(h)$ peut avoir une croissance aussi rapide qu'on veut.

Plaçons-nous maintenant dans le cas de divergence. Chacune des probabilités A_k étant nulle, on peut induire qu'il en est de même de la somme S, et que par suite A_∞ est égal à l'unité. Le résultat est exact, mais le raisonnement précédent manque de rigueur, pour des raisons déjà indiquées. Il est clair, d'autre part, qu'on ne peut ici rechercher la probabilité pour que le cas favorable se produise une infinité de fois sur n épreuves, et faire croître ensuite n indéfiniment; on raisonnera donc comme il suit : Choisissant un nombre fixe m, on cherchera la probabilité pour que le cas favorable se produise plus de m fois sur n épreuves et l'on calculera la limite vers laquelle tend cette probabilité lorsque n augmente indéfiniment; j'omets le calcul facile dont voici le résultat : Cette limite est l'unité quel que soit le nombre fixe m; cela signifie qu'on peut avec avantage parier une somme aussi grande qu'on veut contre 1fr, que le nombre des cas favorables sera supérieur à un nombre fixe quelconque m; c'est précisément la signification de cet énoncé : *la probabilité* A_∞ *est égale à un*. Il est inutile, en effet, de répéter que, dans le cas des probabilités continues, on ne doit pas confondre *probabilité égale à un* avec *certitude*, pas plus qu'on ne doit confondre *probabilité égale à zéro* avec *impossibilité*.

On peut résumer ainsi les résultats obtenus dans l'étude de cette première catégorie de problèmes :

Dans les cas où la série (1) *est convergente, les probabilités* A_0, A_1, ... *ont des valeurs déterminées non nulles et la probabilité* A_∞ *est nulle; dans le cas de la divergence, les probabilités* A_k *sont toutes nulles et* A_∞ *est égale à l'unité.*

5. Passons maintenant à la *seconde catégorie* de problèmes, où les cas possibles sont en infinité dénombrable; supposons d'abord qu'il y ait une seule épreuve et soient $p_1, p_2, p_3, \ldots, p_n, \ldots$ les probabilités des divers cas possibles; la série à termes positifs

$$p_1 + p_2 + \ldots + p_n + \ldots$$

doit, non seulement être convergente, mais encore avoir pour somme l'unité; cette dernière affirmation exige quelques éclaircissements. On pourrait, en effet, se demander si, dans certains cas, une infinité de probabilités nulles ne fournirait pas une somme finie; je me contenterai de dire à ce sujet qu'une telle hypothèse ne me paraît pas logiquement absurde, mais que je n'ai pas rencontré de circonstances où il m'ait paru avantageux de l'introduire. Je me contenterai donc de l'écarter, par une convention expresse, c'est-à-dire que les cas où elle se produirait, s'il en existe, sont écartés des considérations qui suivent; pour pouvoir appliquer nos résultats à un problème particulier, il suffira de constater qu'on n'est effectivement pas dans l'un de ces cas.

Ce qui fait la difficulté et en même temps l'intérêt des problèmes de la seconde catégorie, c'est que les probabilités p_n sont rarement connues avec précision; parfois même, on n'aura sur elles que de vagues renseignements; on devra alors, suivant une méthode indiquée par M. Poincaré pour les probabilités continues, rechercher quelles conclusions générales on peut obtenir, avec le minimum d'hypothèses sur les p_n.

PROBLÈME IV. — *Quelle est la probabilité pour que m épreuves successives donnent des résultats tous différents?*

Désignons par T_m la probabilité demandée; l'emploi de formules classiques conduit aisément à la formule

$$T_m = m!\,\Sigma\, p_{n_1} p_{n_2} \ldots p_{n_m}, \tag{6}$$

le signe Σ s'étendant à toutes les combinaisons de m indices *différents*, chaque combinaison étant prise *une fois seulement.*

Pour évaluer les T_m nous introduisons la fonction entière

$$F(z) = (1+p_1 z)(1+p_2 z)\ldots(1+p_n z)\ldots, \tag{7}$$

qui est de genre zéro, puisque la série à termes positifs Σp_n est convergente. On a visiblement

$$F(z) = 1 + z\Sigma p_1 + z^2 \Sigma p_1 p_2 + \ldots + z^m \Sigma p_{n_1} p_{n_2} \ldots p_{n_m} + \ldots,$$

c'est-à-dire, d'après (6),

$$F(z) = 1 + z + \frac{T_2}{1.2} z^2 + \ldots + \frac{T_m}{m!} z^m + \ldots. \tag{8}$$

La détermination asymptotique de T_m, pour de grandes valeurs de m, dépend de l'*ordre* de la fonction entière $F(z)$; cet ordre, d'après (7), dépend lui-même de l'ordre infinitésimal de p_n. Supposons, pour fixer les idées, qu'on ait, c étant une constante [1],

$$p_n = \frac{c}{n^2};$$

la fonction $F(z)$ est alors d'ordre $\frac{1}{2}$, c'est-à-dire que sa croissance est comparable, à une première approximation dont nous nous contenterons, à celle de la fonction $\cos\sqrt{z}$; les coefficients des deux fonctions sont donc comparables à la même approximation [2]; or, on a

$$\cos\sqrt{z} = \sum \frac{z^m}{(2m)!}.$$

---

[1] La valeur de c se déduit immédiatement de la relation $\Sigma p_n = 1$.

[2] Le fait que tous les p_n sont positifs dans la formule (7) entraîne la conséquence que les coefficients sont tous asymptotiquement égaux à la valeur calculée; sinon, certains d'entre eux pourraient être inférieurs à cette valeur ou même être nuls.

On a donc, *asymptotiquement*,

$$\frac{T_m}{m!} = \frac{1}{(2m)!},$$

c'est-à-dire

$$T_m = \frac{m^m e^{-m}\sqrt{2\pi m}}{2^{2m} m^{2m} e^{-2m}\sqrt{4\pi m}} = \frac{1}{\sqrt{2}}\left(\frac{e}{4m}\right)^m.$$

PROBLÈME V. — *Quelle est la probabilité pour que m épreuves donnent précisément $m-1$ résultats différents?*

Si l'on désigne par V_m cette probabilité, on a

$$V_m = (m-1)!\,\Sigma p_{n_1}^2 p_{n_2} \ldots p_{n_{m-1}}.$$

La méthode classique de Waring donne

$$\Sigma p_{n_1} \Sigma p_{n_2} p_{n_3} \ldots p_{n_m} = m\,\Sigma p_{n_1} p_{n_2} \ldots p_{n_m} + \Sigma p_{n_1}^2 p_{n_2} \ldots p_{n_{m-1}},$$

c'est-à-dire, puisque $\Sigma p_n = 1$,

$$\frac{V_m}{(m-1)!} = \frac{T_{m-1}}{(m-1)!} - m\frac{T_m}{m!},$$

ou enfin

$$V_m = T_{m-1} - T_m.$$

On résoudrait de la même manière les questions analogues.

On voit quelle est l'importance jouée dans ces questions par la manière dont se comporte asymptotiquement p_n, manière qui est caractérisée le plus simplement possible par la fonction $F(z)$. Cette fonction entière interviendra de même dans toutes les questions *où interviendra réellement le fait que les cas possibles* sont *en infinité dénombrable*. On pourrait imaginer en effet des questions, sur lesquelles je juge inutile d'insister, dans l'étude desquelles pourraient être négligés, du moment qu'on a fixé l'approximation désirée dans les calculs, tous les p_n dont l'indice dépasse un certain nombre assignable, toujours le même; tout se passerait dès lors comme si les cas possibles étaient en nombre fini, c'est-à-dire qu'on retomberait dans la théorie classique des probabilités discontinues.

Le caractère essentiel d'un problème de la catégorie qui nous occupe est donc l'ordre infinitésimal de p_n; cet ordre doit être tel que la série à termes positifs Σp_n soit convergente, mais il n'est assujetti *a priori* à aucune condition; étant donnée une série convergente quelconque à termes positifs

$$v = \Sigma v_n,$$

on peut, en effet, poser

$$p_n = \frac{v_n}{v}$$

et l'on a

$$\Sigma p_n = 1.$$

On voit par là combien est grande la variété possible des hypothèses sur les p_n : dans les applications, on sera naturellement conduit à considérer tout d'abord les plus simples, correspondant aux caractères de convergence usuels des séries à termes positifs.

6. Abordons maintenant les questions de la *troisième catégorie*, où l'on suppose à la fois en infinité dénombrable les cas possibles et les épreuves. La probabilité du cas de rang n dans l'épreuve de rang s sera désignée par $p_{n,s}$.

Quel que soit s, la série

$$p_{1,s}+p_{2,s}+\ldots+p_{n,s}+\ldots$$

est convergente et a pour somme l'unité : nous ne répéterons pas ce que nous avons déjà dit à ce sujet pour les problèmes de seconde catégorie. On a ainsi les relations

$$(9)\qquad \begin{cases} p_{1,1}+p_{2,1}+\ldots+p_{n,1}+\ldots=1, \\ p_{1,2}+p_{2,2}+\ldots+p_{n,2}+\ldots=1, \\ p_{1,3}+p_{2,3}+\ldots+p_{n,3}+\ldots=1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{cases}$$

Ces relations sont compatibles avec des hypothèses fort diverses sur les séries telles que la suivante

$$p_{n,1}+p_{n,2}+\ldots+p_{n,s}+\ldots.$$

Si ces dernières séries sont convergentes, quel que soit n, on dira qu'on est dans le *cas de convergence;* on sera dans le *cas de divergence* si elles sont toutes divergentes et dans le *cas mixte* si les unes sont convergentes et les autres divergentes.

Je n'étudierai ici que le cas de convergence; nous poserons

$$(10)\qquad c_n=p_{n,1}+p_{n,2}+\ldots+p_{n,s}+\ldots.$$

On peut observer que la série à termes positifs

$$\Sigma c_n$$

est sûrement divergente, car, si elle était convergente, il en serait de même de la série double

$$\Sigma\Sigma p_{n,s},$$

ce qui est contradictoire avec les relations (9) d'après lesquelles on peut choisir dans cette série des termes dont la somme dépasse tout nombre assignable.

Pour chacun des cas possibles, nous pouvons résoudre les problèmes I, II et III traités plus haut : il suffit de considérer l'arrivée du cas considéré

comme cas favorable et l'arrivée de toute autre comme défavorable. Nous désignerons, pour le cas possible de rang n, par

$$B_{n,0}, \quad B_{n,1}, \quad \ldots, \quad B_{n,k}, \quad \ldots, \quad B_{n,\infty}$$

les probabilités respectivement désignées plus haut par

$$A_0, \quad A_1, \quad \ldots, \quad A_k, \quad \ldots, \quad A_\infty.$$

On désigne par $B_{n,k}$, par exemple, la probabilité pour que le cas possible de rang n se produise précisément k fois dans l'ensemble des épreuves.

Si l'on pose

$$v_{n,s} = \frac{p_{n,s}}{1 - p_{n,s}}, \tag{11}$$

on a

$$\left\{\begin{aligned} B_{n,0} &= \Pi(1 - p_{n,s}), \\ B_{n,1} &= B_{n,0}\,\Sigma\, v_{n,s}, \\ B_{n,k} &= B_{n,0}\,\Sigma\, v_{n,s_1} v_{n,s_2} \ldots v_{n,s_k}, \\ B_{n,\infty} &= 0. \end{aligned}\right. \tag{12}$$

Problème VI. — *Quelle est la probabilité pour que chacun des cas possibles ne se produise qu'un nombre fini de fois?*

Une solution immédiate est fournie par le raisonnement suivant : On a $B_{n,\infty} = 0$; la probabilité pour que le cas de rang n ne se produise qu'un nombre fini de fois est donc égale à l'unité; la probabilité pour qu'il en soit de même quel que soit n est une probabilité composée, égale par conséquent au produit d'une infinité de facteurs tous égaux à l'unité; la valeur de ce produit est l'unité : telle est la probabilité demandée.

Il ne sera pas inutile d'examiner d'un peu plus près la question, afin de se rendre compte de la rigueur du raisonnement : le fait que les facteurs égaux à l'unité sont en infinité dénombrable pourrait, en effet, laisser subsister quelque doute sur la valeur de leur produit.

Cherchons d'abord la probabilité pour que chacun des cas possibles ne se produise pas plus d'une fois. La probabilité, pour le cas de rang n, de ne se produire qu'une fois au plus, est égale à la somme de la probabilité pour qu'il ne se produise pas et de la probabilité pour qu'il se produise précisément une fois, c'est-à-dire à

$$B_{n,0} + B_{n,1}.$$

Les formules (11) et (12) donnent, par des transformations simples,

$$B_{n,0} + B_{n,1} = (1 + v_{n,1} + v_{n,2} + \ldots + v_{n,s} + \ldots) \prod_{s=1}^{s=\infty} \frac{1}{1 + v_{n,s}}.$$

La probabilité pour que chacun des m premiers cas se produise au plus

une fois est donc

$$(13) \qquad \prod_{n=1}^{n=m}(B_{n,0}+B_{n,1}) = \prod_{n=1}^{n=m}\left[\left(1+\sum_{s=1}^{s=\infty} v_{n,s}\right)\prod_{s=1}^{s=\infty}\frac{1}{1+v_{n,s}}\right].$$

Il s'agit de déterminer la limite vers laquelle tend le second membre lorsque m augmente indéfiniment. Or, la série double à termes positifs

$$\Sigma\Sigma\, p_{n,s}$$

étant divergente, il en est de même de la série double à termes positifs

$$\Sigma\Sigma\, v_{n,s}.$$

On en conclut aisément que le produit (13) tend vers zéro lorsque m augmente indéfiniment. En effet, le rapport

$$\frac{\Sigma u_n}{\Pi(1+u_n)}$$

tend vers zéro lorsque les nombres positifs u_n deviennent de plus en plus nombreux et tels que leur somme augmente indéfiniment. Dans les mêmes conditions, le rapport

$$\frac{\Sigma u_n + \Sigma u_{n_1} u_{n_2} + \ldots + \Sigma u_{n_1} u_{n_2} \ldots u_{n_k}}{\Pi(1+u_n)}$$

tend aussi vers zéro quel que soit le nombre fini k; on en conclut que le produit

$$\prod_{n=1}^{n=m}(B_{n,0}+B_{n,1}+\ldots+B_{n,k})$$

tend vers zéro, quel que soit le nombre fixe k, lorsque m augmente indéfiniment. La limite de ce produit pour m infini est égale à *la probabilité pour que chacun des cas possibles se produise au plus k fois*. Cette probabilité est donc nulle, quel que soit le nombre fixe k. Mais il ne faut pas la confondre avec la probabilité pour que chacun des cas possibles ne se produise qu'un nombre fini de fois; car, les cas possibles étant en infinité dénombrable, il peut arriver que le cas de rang n se produise un nombre fini de fois, k_n, mais que, parmi les nombres k_n, il y en ait qui dépassent tout nombre assignable. Nous allons voir que cette éventualité a précisément une probabilité égale à l'unité, les deux autres éventualités possibles, à savoir : 1° *les k_n sont tous inférieurs à un nombre fixe k*; 2° *certains des k_n sont infinis*, ayant toutes deux une probabilité égale à zéro. Nous venons de le voir pour la première de ces éventualités; nous allons le prouver pour la seconde.

Dans ce but, posons

$$S_{n,k} = B_{n,0} + B_{n,1} + \ldots + B_{n,k},$$
$$R_{n,k} = B_{n,k+1} + B_{n,k+2} + \ldots.$$

Nous savons qu'on a

$$S_{n,k} + R_{n,k} = \sum_{s=0}^{s=\infty} B_{n,s} = 1.$$

Étant donné un nombre ε aussi petit qu'on veut, il est clair qu'on peut choisir des nombres k_n tels qu'on ait

$$\sum_{n=1}^{\infty} R_{n,k_n} < \varepsilon,$$

la série du premier membre étant, bien entendu, convergente.

Calculons la probabilité pour que le cas possible de rang n ne se produise pas plus de k_n fois; cette probabilité est

$$\prod_{n=1}^{\infty} S_{n,k_n} = \prod_{n=1}^{\infty} (1 - R_{n,k_n}).$$

Or on a, pour α suffisamment petit $\left(\text{pour } \alpha < \frac{1}{2}\right)$,

$$1 - \alpha > e^{-2\alpha} > 1 - 2\alpha.$$

On en conclut :

$$\prod_{n=1}^{\infty} (1 - R_{n,k_n}) > e^{-2\Sigma R_{n,k_n}} > e^{-2\varepsilon} > 1 - 2\varepsilon.$$

On peut donc, *étant donné le nombre arbitraire* ε, *choisir les nombres* k_n *de manière que la probabilité pour que le cas de rang* n *ne se produise pas plus de* k_n *fois diffère de l'unité de moins de* 2ε.

La probabilité pour que l'un des cas possibles se produise une infinité dénombrable de fois est donc inférieure à 2ε, quel que soit ε, c'est-à-dire qu'elle est égale à zéro.

La détermination effective des k_n, en fonction de ε, dépend de la rapidité de convergence des séries

$$\sum_{s=0}^{\infty} B_{n,s},$$

laquelle dépend elle-même de la rapidité de convergence des séries (com-

parables l'une à l'autre)

$$\sum_{s=1}^{\infty} v_{n,s},$$

$$\sum_{s=1}^{\infty} p_{n,s}.$$

La nature de la convergence de ces dernières séries, lorsque n varie, est la caractéristique essentielle des problèmes de probabilités de la troisième catégorie; leur variété est donc très grande, mais je me bornerai pour le moment à ces indications, ne voulant pas entrer dans des développements théoriques sans applications.

J'espère avoir suffisamment montré, par les quelques problèmes traités d'après quelles méthodes et au moyen de quels principes doivent être abordées les diverses catégories de problèmes de probabilités dénombrables; je vais maintenant faire voir quelle peut être l'utilité des notions nouvelles que nous avons introduites.

II. — Les fractions décimales.

7. Nous considérerons des fractions décimales comprises entre o et 1, c'est-à-dire des expressions de la forme

$$\sum_{n=1}^{\infty} \frac{a_n}{10^n},$$

dans lesquelles les a_n sont des nombres entiers inférieurs à 10. Nous donnerons aussi le nom de *fractions décimales de base q* aux expressions de la forme

$$\sum_{n=1}^{\infty} \frac{b_n}{q^n},$$

dans lesquelles les b_n sont des entiers inférieurs à q : l'emploi de l'épithète *décimal* ne peut prêter à aucune confusion.

8. Nous nous proposons d'étudier la probabilité pour qu'une fraction décimale appartienne à un ensemble donné, en supposant que :

1° Les chiffres décimaux sont indépendants;

2° Chacun d'eux a une probabilité égale à $\frac{1}{q}$ (dans le cas de la base q) de prendre chacun des valeurs possibles : 0, 1, 2, 3, ..., $q-1$.

Il n'est pas besoin d'insister sur le caractère partiellement arbitraire de

ces deux hypothèses; la première, en particulier, est nécessairement inexacte si l'on considère, *comme on est toujours forcé de le faire dans la pratique*, un nombre décimal défini par *une loi*, quelle que soit d'ailleurs la nature de cette loi. Il peut néanmoins être intéressant d'étudier les conséquences de cette hypothèse, afin précisément de se rendre compte de la mesure dans laquelle les choses se passent *comme si* cette hypothèse était vérifiée. La seconde hypothèse, à savoir l'égalité des probabilités pour les diverses valeurs possibles de chaque chiffre décimal, paraît assez naturelle lorsque la première est admise.

L'ensemble des deux hypothèses se justifie d'ailleurs aisément lorsqu'on se place, non au point de vue logique, mais au point de vue géométrique : elles sont, en effet, équivalentes à la suivante : *Le nombre décimal étant représenté par un point du segment* $0-1$, *la probabilité pour qu'il se trouve sur un segment partiel est égale à la longueur de ce segment.* On pourrait interpréter et vérifier à ce point de vue géométrique les résultats que nous allons obtenir; je n'y insisterai pas, afin de laisser entièrement de côté pour le moment la théorie des probabilités continues, qui se rattache, comme je l'ai montré ailleurs, à la théorie de la mesure des ensembles.

9. Si l'on porte son attention sur un chiffre déterminé, par exemple le chiffre 3, et si l'on regarde comme cas favorables ceux où se présente ce chiffre, on est en présence d'un problème de la *première catégorie*, les décimales successives correspondant à l'infinité dénombrable des épreuves. La probabilité du cas favorable étant ici la même à chaque épreuve, nous nous trouvons dans le *cas de divergence;* la probabilité pour que le chiffre 3 soit répété une infinité de fois est donc égale à l'unité ; la probabilité pour que tous les chiffres soient égaux à 3 est égale à zéro; et cependant c'est ce qui arrive si l'on convertit la fraction $\frac{1}{3}$ en fraction décimale; comme nous l'avons observé, *probabilité égale à zéro* n'est pas l'équivalent d'*impossibilité*.

10. Nous allons considérer tout d'abord, pour abréger les écritures, le système de base 2; les seules valeurs possibles des chiffres décimaux sont alors 0 et 1, chacune d'elles ayant la probabilité $\frac{1}{2}$.

Nous appellerons, pour abréger, *cas favorables* les cas où se présente le chiffre 0.

On sait que si l'on considère $2n$ épreuves, la probabilité pour que le nombre des cas favorables soit compris entre

$$n-\lambda\sqrt{n} \quad \text{et} \quad n+\lambda\sqrt{n}$$

est égale à $\theta(\lambda)$, en posant

$$\theta(\lambda)=\frac{2}{\sqrt{\pi}}\int_0^\lambda e^{-\lambda^2}\,d\lambda.$$

Cette probabilité tend très rapidement vers l'unité lorsque λ augmente.

Considérons une suite de nombres λ_n augmentant indéfiniment avec n, mais de telle manière qu'on ait

$$\lim_{n=\infty} \frac{\lambda_n}{\sqrt{n}} = 0.$$

n pourra supposer, par exemple, pour fixer les idées, qu'on prend

$$\lambda_n = \log n.$$

Nous regarderons l'ensemble des $2n$ premières épreuves comme donnant un résultat *favorable* si le nombre de cas où s'est présenté le chiffre 0 est compris entre

$$n - \lambda_n \sqrt{n} \quad \text{et} \quad n + \lambda_n \sqrt{n},$$

et comme donnant un résultat *défavorable* dans le cas contraire. La probabilité p_n du cas favorable est

$$p_n = \Theta(\lambda_n) = \frac{2}{\sqrt{\pi}} \int_0^{\lambda_n} e^{-\lambda^2}\, d\lambda.$$

Si nous considérons le problème de probabilités dénombrables défini par les nombres p_n, on voit que, en posant

$$q_n = 1 - p_n,$$

la série

$$\Sigma\, q_n$$

est convergente. On est donc dans le *cas de convergence* lorsqu'on porte son attention sur les cas défavorables, dont les probabilités sont q_n. La probabilité pour que le cas défavorable se présente une infinité de fois est donc égale à zéro. En d'autres termes, il y a une probabilité égale à *un*, pour que, à partir d'une certaine valeur de n, on se trouve constamment dans le cas favorable. Or, dans ce cas, le rapport entre le nombre des 0 et le nombre des 1 est compris entre

$$\frac{n - \lambda_n \sqrt{n}}{n + \lambda_n \sqrt{n}} \quad \text{et} \quad \frac{n + \lambda_n \sqrt{n}}{n - \lambda_n \sqrt{n}},$$

c'est-à-dire entre

$$\frac{1 - \dfrac{\lambda_n}{\sqrt{n}}}{1 + \dfrac{\lambda_n}{\sqrt{n}}} \quad \text{et} \quad \frac{1 + \dfrac{\lambda_n}{\sqrt{n}}}{1 - \dfrac{\lambda_n}{\sqrt{n}}}.$$

Ce rapport tend donc vers l'unité lorsque n augmente indéfiniment. Donc :

La probabilité pour que le rapport entre le nombre des 0 *et des* 1 *tende vers l'unité* (lorsque le nombre des chiffres considérés augmente indéfiniment) *est égale à un. La probabilité pour que ce rapport ne*

tende vers aucune limite, ou tende vers une limite différente de l'unité, est, par suite, égale à zéro.

11. Les résultats qui viennent d'être énoncés s'étendent sans difficulté aux nombres décimaux de base quelconque; voici, par exemple, la forme qu'ils prennent pour la base 10.

Nous appellerons *fréquence* d'un chiffre décimal jusqu'au rang n le quotient par n du nombre de fois que ce chiffre figure dans les n premières décimales; si la fréquence ainsi définie tend vers une limite lorsque n augmente indéfiniment, on dira que la *fréquence totale* existe et que sa valeur est égale à cette limite.

Ces définitions posées, on a l'énoncé suivant :

La probabilité *pour que la fréquence totale d'un chiffre déterminé existe et soit égale à* $\frac{1}{10}$ a pour valeur l'unité. La probabilité *pour que la fréquence des dix chiffres existe et soit égale à* $\frac{1}{10}$ *pour chacun d'eux* a aussi pour valeur l'unité.

Lorsque la condition précédente est réalisée par une fraction décimale, nous dirons que le nombre égal à cette fraction est *simplement normal* par rapport à la base 10. Un nombre simplement normal est donc caractérisé par le fait que $c_0, c_1, c_2, \ldots, c_8, c_9$ désignant les nombres respectifs de fois que figurent les chiffres 0, 1, 2, ..., 8, 9 parmi les n premières décimales, chacun des rapports

$$\frac{c_0}{n}, \quad \frac{c_1}{n}, \quad \ldots, \quad \frac{c_8}{n}, \quad \frac{c_9}{n}$$

a pour limite $\frac{1}{10}$ lorsque n augmente indéfiniment.

On peut, avec cette définition, dire brièvement : *La probabilité pour qu'un nombre soit simplement normal par rapport à la base* 10 *est égale à l'unité.*

12. Tout nombre écrit dans une base déterminée, 10, par exemple, peut être, sans aucun calcul, regardé comme écrit dans une base égale à une puissance quelconque de 10.

Nous dirons qu'un nombre est *entièrement normal* par rapport à la base 10 (ou, pour abréger, *normal* par rapport à cette base), lorsque ce nombre et ses produits par les diverses puissances de 10 sont *simplement normaux* par rapport à toutes les bases égales à une puissance de 10 :

$$10, \quad 100, \quad 1000, \quad \ldots, \quad 10^n, \quad \ldots.$$

La probabilité pour que cette condition soit remplie est égale à l'unité pour chacune de ces bases; la probabilité pour qu'un nombre soit entiè-

rement normal par rapport à la base 10 est donc aussi égale à l'unité (¹).

La propriété caractéristique d'un nombre normal est la suivante : *Un groupement quelconque de p chiffres consécutifs étant considéré, si l'on désigne par c_n le nombre de fois que se rencontre ce groupement dans les n premiers chiffres décimaux, on a*

$$\lim_{n=\infty} \frac{c_n}{n} = \frac{1}{10^p}.$$

13. Lorsqu'un nombre donné est normal *par rapport à toutes les bases possibles*, nous dirons qu'il est *absolument normal*. La probabilité pour qu'un nombre soit absolument normal est encore égale à l'unité. Dans l'état actuel de la science, la détermination effective d'un nombre absolument normal paraît un problème des plus difficiles : il serait intéressant, soit de le résoudre en construisant un nombre absolument normal, ou en montrant qu'un nombre irrationnel connu est absolument normal, soit de démontrer que, parmi les nombres pouvant être réellement définis, aucun n'est absolument normal; si paradoxal que paraisse cet énoncé, il n'est nullement incompatible avec le fait que la probabilité pour qu'un nombre soit absolument normal est égale à l'unité (²).

14. On peut aisément construire un nombre normal par rapport à la base 10 ; quelques précautions sont cependant nécessaires si l'on veut définir sans ambiguïté un nombre unique et déterminé. Voici comment on peut procéder.

Considérons l'ensemble des nombres entiers de n chiffres au plus et écrivons, à la gauche de ceux qui ont moins de n chiffres, des zéros en nombre suffisant pour qu'ils aient précisément n chiffres : nous dirons qu'un tel nombre de n chiffres est de catégorie p lorsque la différence entre le nombre de fois qu'y figure le chiffre qui y figure le plus souvent et le nombre de fois qu'y figure le chiffre qui y figure le moins souvent est égale à p. Il n'y a des nombres de catégorie zéro que si n est un multiple de 10; tel est le nombre : 1234567890; dans ce cas, il n'y a pas de nombre de catégorie 1 ; si n n'est pas un multiple de 10, il n'y a pas de nombre de catégorie 0, mais il y en a de catégorie 1; par exemple les nombres 145 ou 1234567890123; quel que soit n, il y a des nombres appartenant aux catégories 2, 3, ..., $n-1$, n. On formera un groupe de chiffres G_n de la manière suivante : on écrira les uns à la suite des autres les nombres de n chiffres de catégorie 0 ou 1, en les rangeant par ordre de grandeur croissante; ensuite, les nombres de n chiffres de catégorie 2, toujours par ordre de grandeur croissante, puis ceux de catégorie 3 et ainsi de suite, jusqu'à ceux de catégorie n.

(¹) Je juge inutile de revenir sur la démonstration détaillée du fait qu'on a bien le droit d'appliquer le théorème des probabilités composées, malgré l'infinité dénombrable des cas.

(²) Depuis que ceci a été écrit, M. Lebesgue m'a indiqué une méthode pour définir un nombre absolument normal.

On obtient un nombre décimal normal en écrivant après la virgule, successivement les groupes G_1, G_2, ..., G_n, ... Je me bornerai à indiquer que le point essentiel, dans la démonstration de ce fait, est le suivant : Le nombre total des chiffres des nombres de n chiffres et de catégorie o ou 1 est, pour n infiniment grand, infiniment petit par rapport au nombre total des chiffres de tous les nombres de $n-1$ chiffres.

15. Au lieu de considérer l'ensemble de tous les nombres décimaux, bornons-nous maintenant aux fractions décimales périodiques. Si l'on convient de ne pas écrire de fraction décimale limitée, c'est-à-dire de remplacer 0,43, par exemple, par la fraction périodique équivalente 0,42999..., l'ensemble des fractions décimales périodiques est identique à l'ensemble des nombres rationnels. Peut-on définir la probabilité pour que la partie non périodique ait précisément i chiffres? pour que la période ait précisément k chiffres? pour que la partie non périodique ou la période soient formées par tels chiffres fixés d'avance? Nous nous trouvons ici en présence d'un problème de seconde catégorie : les cas possibles sont en infinité dénombrable. Si ces cas étaient regardés tous comme également probables, la probabilité de chacun ne pourrait pas avoir une valeur finie quelconque; on serait conduit à la regarder comme égale à zéro et à supposer que l'ensemble de ces probabilités nulles a une somme égale à l'unité; c'est là une conception inacceptable; l'hypothèse que la probabilité pour que la période ait un seul chiffre, par exemple, a pour valeur zéro, n'est pas moins inacceptable; quel que soit le procédé adopté pour définir une fraction périodique arbitraire [1], il est certain que la probabilité pour que la période ait un seul chiffre ne peut pas être regardée comme nulle; car ce fait se produira certainement un certain nombre de fois sur un nombre fini d'expériences.

Nous désignerons par $p_{i,k}$ la probabilité pour que la partie non périodique se compose de i chiffres et que la période ait k chiffres. Le nombre 1 peut avoir une valeur entière quelconque, y compris zéro, et le nombre k une valeur entière quelconque, zéro exclu. On a nécessairement

$$\sum_{i=0}^{\infty} \sum_{k=1}^{\infty} p_{i,k} = 1. \tag{14}$$

Nous admettrons que les divers chiffres 0, 1, 2, ..., 9 ont des probabilités égales; la probabilité d'un nombre *déterminé* dont la partie non périodique se compose de i chiffres et la période de k chiffres est, dans cette hypothèse, si l'on suppose que k est un nombre premier,

$$\frac{p_{i,k}}{10^i(10^k-10)}, \tag{15}$$

[1] Il faut, bien entendu, que le procédé soit tel qu'il conduise *sûrement* à une fraction *périodique*: le procédé qui consisterait à imaginer qu'on tire au sort les chiffres successifs n'est pas admissible.

car il y a pour la partie non périodique 10^i possibilités également probables, et pour la partie périodique $10^k - 10$ seulement, car on doit exclure les nombres de k chiffres formés de chiffres identiques. Si le nombre k n'était pas premier, l'expression serait plus compliquée : il est inutile de l'écrire; observons simplement que son ordre de grandeur n'est pas sensiblement modifié, c'est-à-dire que le facteur correctif diffère peu de l'unité et tend vers l'unité lorsque k augmente indéfiniment. Par exemple, pour $k = 30$, le diviseur $10^k - 10$ doit être remplacé par

$$10^{30} - 10^{15} - 10^{10} - 10^6 + 10^5 + 10^3 + 10^2 - 10.$$

16. Les formules (14) et (15) sont essentielles; notons aussi les suivantes : on posera

$$\sum_{k=1}^{\infty} p_{i,k} = P_i \qquad (i = 0, 1, 2, \ldots),$$

$$\sum_{i=0}^{\infty} p_{i,k} = Q_k \qquad (k = 1, 2, \ldots),$$

et l'on aura

$$\sum_{i=0}^{\infty} P_i = \sum_{k=1}^{\infty} Q_k = 1.$$

On définira ensuite les fonctions entières

$$F(z) = \prod_{i=0}^{\infty} (1 + P_i z),$$

$$G(z) = \prod_{k=1}^{\infty} (1 + Q_k z),$$

$$\varphi(z) = \prod_{i=0}^{\infty} \prod_{k=1}^{\infty} (1 + p_{i,k} z)$$

et il sera nécessaire de faire des hypothèses sur l'ordre de ces fonctions entières.

Un cas particulier intéressant serait celui où l'on supposerait qu'on peut poser

$$p_{i,k} = p_i q_k.$$

Il y aurait alors avantage à considérer séparément la convergence des séries

$$\sum_{i=0}^{\infty} p_i = 1,$$

$$\sum_{h=1}^{\infty} q_k = 1$$

ou, ce qui revient au même, les fonctions entières

$$f(z) = \prod_{i=0}^{\infty} (1 + p_i z),$$

$$g(z) = \prod_{k=1}^{\infty} (1 + q_k z).$$

Je me bornerai à ces indications, car nous allons rencontrer, en étudiant les fractions continues, des problèmes analogues, mais qui me paraissent plus intéressants.

III. — Les fractions continues.

17. Considérons le développement en fraction continue d'un nombre irrationnel compris entre 0 et 1 ; c'est un développement de la forme

$$x = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ldots}}}, \tag{16}$$

dans lequel les *quotients incomplets* $a_1, a_2, \ldots, a_n, \ldots$ sont des nombres entiers pouvant prendre toutes les valeurs. L'étude des problèmes de probabilités qu'on peut se poser à propos d'un tel développement rentre donc dans notre *troisième catégorie*, puisque nous avons une infinité dénombrable de nombres a_n dont chacun peut acquérir une infinité dénombrable de valeurs. Nous désignerons, d'une manière générale, par $p_{i,k}$ la probabilité pour que le quotient incomplet a_i soit égal au nombre entier k. Les nombres entiers i et k peuvent prendre toutes les valeurs entières différentes de zéro (1); et l'on a, quel que soit i, la relation

$$\sum_{k=1}^{\infty} p_{i,k} = 1.$$

On pourrait, *a priori*, faire sur les $p_{i,k}$ des hypothèses arbitraires, pourvu que les relations précédentes soient satisfaites ; nous allons étudier les hypothèses auxquelles on est conduit lorsqu'on se place au point de vue géométrique déjà signalé à propos des nombres décimaux.

18. Calculons d'abord les nombres $p_{i,k}$, c'est-à-dire cherchons la probabilité pour qu'on ait

$$a_1 = k.$$

(1) On pourrait admettre que, si la fraction continue représente un nombre rationnel, le dernier quotient (complet) ayant le rang n, le quotient a_{n+1} est égal à 1 et les suivants sont égaux à zéro ; mais cette convention ne me paraît présenter aucun avantage.

D'après la relation (16), ceci entraîne

$$\frac{1}{k+1} < x < \frac{1}{k}.$$

Le nombre x doit donc se trouver dans un intervalle de longueur

$$\frac{1}{k} - \frac{1}{k+1} = \frac{1}{k(k+1)}.$$

La probabilité géométrique pour que x appartienne à cet intervalle est égale à sa longueur; si donc nous adoptons cette probabilité géométrique, nous aurons

$$p_{1,k} = \frac{1}{k(k+1)}.$$

On a, en particulier,

$$p_1 = \frac{1}{2}, \qquad p_2 = \frac{1}{6}, \qquad p_3 = \frac{1}{12}, \qquad p_4 = \frac{1}{20}, \qquad \dots.$$

Passons maintenant aux $p_{2,k}$. Si $a_1 = n$, la condition pour que a_2 soit égal à k sera, d'après (16),

$$\frac{1}{n+\frac{1}{k}} < x < \frac{1}{n+\frac{1}{k+1}},$$

c'est-à-dire

$$\frac{k}{nk+1} < x < \frac{k+1}{n(k+1)+1}.$$

Le nombre x doit être compris dans un intervalle de longueur

$$\frac{k+1}{n(k+1)+1} - \frac{k}{nk+1} = \frac{1}{(nk+n+1)(nk+1)}.$$

La probabilité $p_{2,k}$ est égale à la somme des longueurs des intervalles analogues, correspondant aux diverses valeurs de n; on a donc

$$p_{2,k} = \sum_{n=1}^{\infty} \frac{1}{(nk+n+1)(nk+1)},$$

ce qu'on peut écrire aussi sous les formes suivantes :

$$p_{2,k} = \sum_{n=1}^{\infty} \frac{1}{n^2}\left(\frac{1}{k+\frac{1}{n}} - \frac{1}{k+1+\frac{1}{n}}\right),$$

$$p_{2,k} = \frac{1}{k(k+1)} \sum_{n=1}^{\infty} \frac{1}{\left(n+\frac{1}{k}\right)\left(n+\frac{1}{k+1}\right)}.$$

Cette dernière forme permet de calculer la valeur asymptotique de $p_{2,k}$ pour de grandes valeurs de k; en effet, pour k infini, la somme qui y figure se réduit à

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

On a donc la relation

$$p_{2,k} = \frac{\pi^2}{6} \frac{1}{k(k+1)}(1 - \varepsilon_k),$$

le nombre positif ε_k tendant vers zéro lorsque k augmente indéfiniment.

On voit que, pour de grandes valeurs de k, $p_{2,k}$ est supérieur à $p_{1,k}$; donc pour de petites valeurs de k, $p_{2,k}$ doit être inférieur à $p_{1,k}$, puisqu'on a

$$\sum_{k=1}^{\infty} p_{1,k} = \sum_{k=1}^{\infty} p_{2,k} = 1.$$

On trouve effectivement, par un calcul facile,

$$p_{2,1} = 2\log 2 - 1 = (4\log 2 - 2)p_{1,1}.$$

Or on a visiblement

$$4\log 2 - 2 < 1,$$

car cette égalité revient à

$$2^4 < e^3.$$

On calculerait de la même manière les probabilités $p_{3,k}$. Si $a_1 = n$ et $a_2 = n'$, la condition pour que a_3 soit égal à k est

$$\cfrac{1}{n + \cfrac{1}{n' + \cfrac{1}{k+1}}} < x < \cfrac{1}{n + \cfrac{1}{n' + \cfrac{1}{k}}}$$

c'est-à-dire

$$\frac{n'(k+1)+1}{(nn'+1)(k+1)+n} < x < \frac{n'k+1}{(nn'+1)k+n}.$$

L'intervalle défini par les inégalités précédentes a pour longueur

$$\frac{1}{[(nn'+1)(k+1)+n][nn'+1)k+n]},$$

de sorte qu'on a

$$p_{3,k} = \frac{1}{k(k+1)} \sum_{n=1}^{\infty} \sum_{n'=1}^{\infty} \frac{1}{\left(nn'+1+\frac{n}{k+1}\right)\left(nn'+1+\frac{n}{k}\right)}.$$

Lorsque k augmente indéfiniment, la somme double tend vers la limite

$$\sum_{n=1}^{\infty}\sum_{n'=1}^{\infty}\frac{1}{(nn'+1)^2}=\sum_{q=2}^{\infty}\frac{\theta(q-1)}{q^2},$$

où l'on a désigné par $\theta(\mathrm{N})$ le nombre des diviseurs de N, y compris l'unité et N lui-même.

On voit que le calcul successif des divers nombres $p_{i,k}$ conduit à des calculs intéressants, mais rapidement compliqués, même pour de petites valeurs de i; il nous suffira, pour l'instant, d'avoir indiqué la marche à suivre dans ces calculs; les nombres $p_{i,k}$ peuvent être regardés comme connus, en ce sens que le calcul effectif de l'un d'eux, avec une approximation donnée à l'avance, peut être effectué en un temps fini par une méthode régulière. Nous allons nous attacher à rechercher des inégalités auxquelles satisfont ces nombres.

19. Désignons par

$$\frac{\mathrm{P}_{n-2}}{\mathrm{Q}_{n-2}}\quad\text{et}\quad\frac{\mathrm{P}_{n-1}}{\mathrm{Q}_{n-1}}$$

les réduites de rang $n-2$ et $n-1$ de la fraction continue; pour qu'on ait $a_n=k$, il faut et il suffit que x soit compris entre

$$\frac{\mathrm{P}_{n-2}+k\mathrm{P}_{n-1}}{\mathrm{Q}_{n-2}+k\mathrm{Q}_{n-1}}\quad\text{et}\quad\frac{\mathrm{P}_{n-2}+(k+1)\mathrm{P}_{n-1}}{\mathrm{Q}_{n-2}+(k+1)\mathrm{Q}_{n-1}}.$$

La différence entre ces deux fractions (dont la grandeur relative dépend de la parité de n) est, en valeur absolue,

$$(17)\qquad \frac{1}{(\mathrm{Q}_{n-2}+k\mathrm{Q}_{n-1})[\mathrm{Q}_{n-2}+(k+1)\mathrm{Q}_{n-1}]} = \frac{1}{\mathrm{Q}_{n-1}^2}\,\frac{1}{\left(k+\dfrac{\mathrm{Q}_{n-2}}{\mathrm{Q}_{n-1}}\right)\left(k+1+\dfrac{\mathrm{Q}_{n-2}}{\mathrm{Q}_{n-1}}\right)}.$$

Telle est la longueur d'un des intervalles dont la somme est égale à $p_{n,k}$. De même $p_{n,k+1}$ est égal à la somme d'intervalles égaux à

$$(18)\qquad \frac{1}{\mathrm{Q}_{n-1}^2}\,\frac{1}{\left(k+1+\dfrac{\mathrm{Q}_{n-2}}{\mathrm{Q}_{n-1}}\right)\left(k+2+\dfrac{\mathrm{Q}_{n-2}}{\mathrm{Q}_{n-1}}\right)}.$$

Le rapport de l'intervalle (18) à l'intervalle correspondant (17) est égal à

$$(19)\qquad \frac{k+\dfrac{\mathrm{Q}_{n-2}}{\mathrm{Q}_{n-1}}}{k+2+\dfrac{\mathrm{Q}_{n-2}}{\mathrm{Q}_{n-1}}}.$$

Le rapport de $p_{n,k+1}$ à $p_{n,k}$ est compris entre la plus grande et la plus petite des valeurs que prend le rapport (19) lorsque Q_{n-2} et Q_{n-1} prennent toutes les valeurs possibles [étant spécifié que Q_{n-2} et Q_{n-1} sont les dénominateurs respectifs de la $(n-2)^{\text{ième}}$ et de la $(n-1)^{\text{ième}}$ réduite d'une même fraction continue]. Or on a [1]

$$0 < \frac{Q_{n-2}}{Q_{n-1}} < 1.$$

On en conclut que le rapport (19) est compris entre $\frac{k}{k+2}$ et $\frac{k+1}{k+3}$; on a donc

$$\frac{k}{k+2} < \frac{p_{n,k-1}}{p_{n,k}} < \frac{k+1}{k+3}.$$

On obtient donc, par récurrence,

$$\frac{(k-1)(k-2)\ldots 2.1}{(k+1)k\ldots 4.3} < \frac{p_{n,k}}{p_{n,1}} < \frac{k(k-1)\ldots 3.2}{(k+2)(k+1)\ldots 5.4},$$

c'est-à-dire

$$\frac{2}{k(k+1)} < \frac{p_{n,k}}{p_{n,1}} < \frac{6}{(k+1)(k+2)}. \tag{20}$$

Utilisons la relation

$$\sum_{k=1}^{\infty} p_{n,k} = 1; \tag{21}$$

nous obtenons

$$\sum_{k=1}^{\infty} \frac{2}{k(k+1)} < \frac{\sum_{1}^{\infty} p_{n,k}}{p_{n,1}} < \sum_{k=1}^{\infty} \frac{6}{(k+1)(k+2)},$$

c'est-à-dire

$$2 < \frac{1}{p_{n,1}} < 3,$$

ou enfin

$$\frac{1}{3} < p_{n,k} < \frac{1}{2}. \tag{22}$$

En multipliant membre à membre les inégalités (20) et (22), on obtient enfin

$$\frac{2}{3k(k+1)} < p_{n,k} < \frac{3}{(k+1)(k+2)}. \tag{23}$$

Il suffira de conserver cette relation (22), car, pour $k = 1$, elle se réduit

---

[1] On pourrait avoir une approximation meilleure en tenant compte des résultats sur la probabilité de Q_{n-1}; l'inégalité employée nous suffira.

à (22); elle nous fait connaître les nombres $p_{n,k}$ avec une précision suffisante à une première approximation : le rapport des deux limites entre lesquelles ils sont enfermés est inférieur à $\frac{9}{2}$.

20. Les inégalités (23) montrent que la série

$$\sum_{n=1}^{\infty} p_{n,k}$$

est divergente quel que soit k; je n'insisterai pas sur les conséquences qu'on peut tirer de ce fait, le cas de convergence étant plus intéressant.

21. Nous poserons

$$P_{n,k} = p_{n,1} + p_{n,2} + \ldots + p_{n,k}.$$

Cette somme $P_{n,k}$ est la probabilité pour que a_n soit *égal ou inférieur à* k; les inégalités (23) permettent d'obtenir des limites entre lesquelles est compris $P_{n,k}$. Mais, pour obtenir des limites approchées, il est avantageux de remarquer qu'on a, en vertu de (21),

$$1 - P_{n,k} = p_{n,k+1} + p_{n,k-2} + \ldots$$

et d'appliquer les inégalités (23) aux divers termes du second membre; on obtient ainsi

$$\frac{2}{3(k+1)} < 1 - P_{n,k} < \frac{3}{k+2},$$

c'est-à-dire

$$1 - \frac{3}{k+2} < P_{n,k} < 1 - \frac{2}{3(k+1)}.$$

Prenons pour k une fonction croissante de n, soit $\varphi(n)$, telle que la série

$$\sum \frac{1}{\varphi(n)} \tag{24}$$

soit convergente. La série simple

$$\sum_{n=1} [1 - P_{n.\varphi(n)}]$$

est alors convergente. On en conclut la proposition suivante :

La fonction $\varphi(n)$ *étant telle que la série* (24) *soit convergente, la probabilité pour qu'on ait, pour une infinité de valeurs de* n,

$$a_n > \varphi(n)$$

est égale a zéro. Autrement dit, *il y a une probabilité égale à un pour qu'on ait, à partir d'une valeur finie de n,*

$$a_n < \varphi(n).$$

Quelle que soit la fonction $\varphi(n)$ telle que la série (24) soit convergente, il existe une fonction $\psi(n)$ satisfaisant à la même condition et telle qu'on ait

$$\lim_{n=\infty} \frac{\psi(n)}{\varphi(n)} = 0.$$

On en conclut que, *si la série (24) est convergente, il y a une probabilité égale à un pour qu'on ait*

$$\lim_{n=\infty} \frac{a_n}{\varphi(n)} = 0. \tag{25}$$

Si la série (24) était divergente, on ne saurait affirmer que la limite précédente est égale a l'infini, mais seulement la relation

$$\overline{\lim_{n=\infty}} \frac{a_n}{\varphi(n)} = \infty, \tag{26}$$

où la notation $\overline{\lim}$ due à M. Pringsheim désigne la *plus grande limite* définie par Cauchy et précisée par Paul du Bois Reymond et par M. Hadamard.

Si l'on ne conserve que ceux des a_n qui interviennent dans la lim, les relations (25) et (26) expriment ce fait fort curieux, que le mode de croissance ainsi défini par les a_n est, *avec une probabilité égale à l'unité*, inférieur à toute fonction $\varphi(n)$ telle que la série (24) soit convergente, et supérieur a toute fonction $\varphi(n)$ telle que la série (24) soit divergente [1]. Ce résultat apporte une contribution nouvelle à la question si difficile de la limite entre la convergence et la divergence. A ce titre, il me paraît être le plus intéressant de ceux que nous avons obtenus dans ce Mémoire.

---

[1] Lorsque l'on considère une suite *déterminée* a_n il est clair que la série

$$\sum \frac{1}{a_n}$$

est ou convergente, ou divergente, de sorte que, suivant les cas, il existe une infinité de fonctions $\varphi(n)$ inférieures a a_n et telles que la série (24) soit convergente, ou une infinité de fonctions $\varphi(n)$ supérieures a a_n et telles que la série (24) soit divergente; mais lorsqu'on *fixe* la fonction $\varphi(n)$, il y a une probabilité égale à zéro pour que la suite a_n, définie par la fraction continue, satisfasse à l'une des conditions précédentes.

IV. — Questions diverses et conclusion.

22. Les applications des probabilités dénombrables me paraissent ne pas devoir se borner au développement des recherches précédentes; je voudrais indiquer rapidement d'autres questions dans lesquelles on les ferait intervenir utilement.

23. Considérons une équation du second degré *à coefficients entiers*,

$$ax^2 - bx + c = 0, \tag{27}$$

et soit

$$\delta = b^2 - 4ac.$$

Pour que l'équation ait ses racines rationnelles, il est nécessaire et suffisant que le nombre entier δ soit carré parfait; si δ est nul, les racines sont égales.

On peut se proposer de déterminer : la probabilité P pour que l'équation (27) ait ses racines rationnelles; la probabilité P′ pour qu'elle ait ses racines égales; la probabilité ϖ pour qu'elle ait ses racines réelles; etc.

Nous désignerons par a_m la probabilité pour que le coefficient a soit égal à l'entier positif ou négatif m; on supposera [1] $a \neq 0$, pour que l'équation (17) soit effectivement du second degré et l'on aura la relation

$$\sum_{-\infty}^{+\infty}{}' a_m = 1, \tag{28}$$

le signe $\sum'$ indiquant que la valeur zéro est exclue. On a, de même,

$$\sum_{-\infty}^{+\infty} b_n = 0, \tag{29}$$

$$\sum_{-\infty}^{+\infty}{}' c_p = 0, \tag{30}$$

les valeurs n et p de b et de c pouvant être des nombres entiers quelconques y compris zéro. Ceci suppose les trois coefficients indépendants : nous nous bornerons à cette hypothèse.

La probabilité P pour que l'équation (27) ait ses racines rationnelles est

---

(1) Cette hypothèse n'est pas indispensable.

évidemment

$$(31) \qquad P = \sum a_m b_n c_p,$$

la sommation étant étendue à toutes les valeurs de m, n, p telles que $n^2 - 4mp$ soit carré parfait. On a de même

$$(32) \qquad P' = \sum a_m b_n c_p \qquad (n^2 - 4mp = 0),$$

$$(33) \qquad \varpi = \sum a_m b_n c_p \qquad (n^2 - 4mp \gtreqless 0).$$

Les séries (31), (32), (33) sont évidemment convergentes, puisque les séries à termes positifs (28), (29), (30) le sont. Pour les évaluer effectivement, il serait nécessaire de préciser les hypothèses sur les valeurs de a_m, b_n, c_p. On pourrait, par exemple, considérer un développement en fraction continue tel que (16) et poser

$$a = (-1)^{a_4} a_1,$$
$$b = (-1)^{a_5}(a_2 - 1),$$
$$c = (-1)^{a_6}(a_3 - 1).$$

Ou encore

$$a = a_1,$$
$$b = a_2 - a_3,$$
$$c = a_3 - a_4.$$

Mais il serait surtout intéressant d'étudier des hypothèses qui s'imposeraient naturellement par l'étude d'un problème concret.

24. Je signale seulement en passant les extensions diverses qu'on pourrait faire aux équations de degré supérieur : on pourra même *laisser indéterminé le degré de l'équation* et rechercher, par exemple, la probabilité pour qu'une équation algébrique arbitraire à coefficients entiers ait une racine rationnelle.

25. Les considérations précédentes conduisent naturellement à poser la question suivante : on considère un corps quelconque de nombres, par exemple l'ensemble des nombres algébriques, et l'on demande la probabilité pour qu'un nombre choisi arbitrairement dans ce corps appartienne à une classe donnée (infinie ou finie, ou même formée d'un seul nombre). Une telle question ne peut évidemment être résolue qu'au moyen d'hypothèses comportant une part d'arbitraire : cette constatation ne diminue pas l'intérêt qui peut s'attacher à son étude; on pourrait même dire sans paradoxe qu'elle l'augmente, en permettant, par la variété des hypothèses, la solution de problèmes plus nombreux.

26. D'une manière plus générale, il est clair que tous les éléments ana-

lytiques, nombres ou fonctions, qui peuvent être effectivement définis, sont en infinité dénombrable; on pourra donc se poser à leur sujet des problèmes analogues aux précédents. Par *effectivement définis* on doit entendre : *définis au moyen d'un nombre fini de mots*, et il est clair que les éléments pour lesquels le nombre de mots nécessaires à la définition est extrêmement grand devront être regardés comme ayant une probabilité extrêmement petite. On peut regarder, par exemple, comme pratiquement nulle la probabilité pour que l'on considère une fonction ou un nombre dont la définition exigerait un million de volumes de cinq cents pages (¹).

27. Lorsque la théorie des probabilités dénombrables aura été développée dans le sens qui vient d'être indiqué, il sera intéressant de comparer les résultats acquis avec ceux qu'on obtient par la théorie des probabilités continues ou géométriques.

Il *existe* certainement (si ce n'est pas un abus d'employer ici le verbe *exister*) dans le continu géométrique des éléments qui ne peuvent pas être définis : tel est le sens réel de l'importante et célèbre proposition de M. Georg Cantor : le continu n'est pas dénombrable. Le jour où ces éléments *indéfinissables* seront réellement mis à part et où l'on ne prétendrait point les faire intervenir plus ou moins implicitement, il en résulterait certainement une grande simplification dans les méthodes de l'Analyse; je serais heureux si les pages précédentes pouvaient contribuer à faire pressenti. l'intérêt qui s'attacherait à l'étude de telles questions.

IV. — Sur un problème de probabilités relatives aux fractions continues (²).

Dans un récent article (³), M. Félix Bernstein croit découvrir une contradiction entre un de ses résultats et un énoncé que j'ai donné il y a quelques années (⁴). Il n'y a en réalité aucune contradiction : il suffit pour s'assurer qu'il ne peut pas y en avoir, de constater l'analogie complète entre les formules ci-dessus pages 205-206 et pages 425-427 (Be); cette analogie est parfois masquée par les différences de notations. De plus, M. Bernstein

(¹) Ceci peut être rapproché de considérations sur la *hauteur* que j'ai développées ailleurs : Borel, *Contribution à l'analyse arithmétique du continu* (*Journal de Mathématiques pures et appl.*, 5ᵉ série, t. IX, 1903, p. 329-375).

(²) *Math. Annalen*, t. LXXII.

(³) Félix Bernstein, *Ueber eine Anwendung der Mengenlehre auf ein aus der Theorie der Säkularen Störungen herrührendes Problem* (*Math. Ann.*, t. LXXI, 1911, p. 417-439). Je désignerai ce travail par Be.

(⁴) Émile Borel, *Les probabilités dénombrables et leurs applications arithmétiques* (*Rend. Circ. Mat. Palermo*, t. XXVII, 1909, p. 247-271). Ce travail constitue les trois premiers paragraphes de cette Note V.

fait sur l'indépendance des probabilités une remarque qui, si elle n'infirme en rien mes résultats, est néanmoins juste en elle-même. Il ne me paraît donc pas inutile de reprendre brièvement la question.

Ce qui est exact dans l'objection de M. Bernstein, c'est que le raisonnement que j'ai donné pages 183-186 suppose les probabilités indépendantes et doit être modifié lorsqu'elles ne le sont pas ([1]). Mais cette modification du raisonnement est aisée et n'entraîne aucune modification dans le résultat, c'est ce que je voudrais montrer ici, en développant complètement le raisonnement dans le cas où les probabilités ne sont pas indépendantes. On constatera que cette nouvelle démonstration ne diffère pas sensiblement de ce que deviendrait la démonstration du Mémoire cité si l'on y développait tous les calculs *sous forme entière*.

Nous considérons une infinité dénombrable d'épreuves successives, qui sont numérotées à l'aide des nombres entiers dans leur ordre naturel; la probabilité du cas favorable est désignée par p_n pour l'épreuve de rang n; la probabilité du cas défavorable est $q_n = 1 - p_n$.

Nous supposons que p_n et q_n désignent les probabilités globales, lorsqu'on ignore le résultat des $n-1$ premières épreuves; lorsqu'on connaît ces résultats, la probabilité, au lieu d'être p_n, aura généralement une valeur différente, comprise cependant par hypothèse entre des limites connues p'_n et p''_n.

$$p'_n \leqq p_n \leqq p''_n.$$

Nous supposerons essentiellement tous les p'_n et tous les p''_n différents de 0 et de 1. Nous dirons que nous sommes dans le cas de convergence si les deux séries

$$\sum p'_n, \tag{1$'$}$$

$$\sum p''_n \tag{1$''$}$$

sont convergentes et si nous sommes dans le cas de divergence lorsque ces deux séries sont divergentes ([2]). Nous allons étudier le cas de convergence ([3]); nous désignerons par A_0 la probabilité pour que le cas favorable ne se produise jamais, par A_k la probabilité pour que ce cas favorable

---

([1]) Je dois dire que cette objection m'avait été faite, dans une lettre particulière, par M. Lebesgue, au moment de la publication des *Rendiconti*. Je m'étais assuré que mes résultats étaient exacts et par suite n'avais pas attaché d'importance à cette objection; j'en avais même perdu le souvenir quand, quelques années plus tard, j'ai répondu à M. Bernstein; c'est seulement après la publication de cette réponse reproduite ici que j'ai retrouvé l'ancienne lettre de M. Lebesgue.

([2]) On voit que, dans le cas où les probabilités ne sont pas indépendantes, il peut arriver que la série (1)' converge et que la série (1)'' diverge; c'est un cas nouveau que je laisse de côté aujourd'hui.

([3]) On verrait d'une manière analogue que les conclusions de mon Mémoire subsistent dans le cas de divergence.

se produise k fois et k fois seulement, par A_∞ la probabilité pour qu'il se produise une infinité de fois. Il est évident qu'on a

$$A_0 + A_1 + A_2 + \ldots - A_\infty = 1.$$

Notre but est de démontrer qu'on a

$$A_\infty = 0.$$

Il suffit donc de prouver la relation

$$A_0 + A_1 + \ldots + A_n + \ldots = 1.$$

Cette série à termes positifs ayant une somme bien déterminée, il suffit évidemment de prouver qu'elle diffère de l'unité aussi peu qu'on veut. Pour cela nous remarquerons que, la série $(1)''$ étant convergente, on peut, à tout nombre donné ε, faire correspondre un nombre k tel qu'on ait

$$\prod_{n=k+1}^{\infty} (1 - p_n'') > 1 - \varepsilon.$$

Nous allons évaluer la somme

$$S_k = A_0 + A_1 + \ldots + A_k$$

et démontrer que cette somme est supérieure à $1 - \varepsilon$; il en résultera que la série à termes positifs, qui ne peut dépasser 1, est bien égale à 1.

Calculons les divers termes de cette somme et, pour cela, précisons les notations.

La probabilité du cas favorable à la première épreuve a été désignée par p_1 et celle du cas défavorable par q_1; nous poserons, pour rendre les notations symétriques,

$$q_1 = p_0.$$

Dans le cas où la première épreuve a été favorable, nous désignerons par p_{11} la probabilité favorable à la deuxième épreuve et par p_{10} la probabilité du cas défavorable; si la première épreuve a été défavorable, la probabilité du cas favorable à la deuxième épreuve sera p_{01} et la probabilité du cas défavorable sera p_{00}.

De même, p_{011} désigne la probabilité du cas favorable à la troisième épreuve lorsque la première a été défavorable et la seconde favorable; et p_{0000} désigne la probabilité du cas défavorable à la quatrième épreuve lorsque les trois premières ont été défavorables, etc.

On a évidemment

$$p_1 p_{11} + p_0 p_{01} = p_2,$$
$$p_1 p_{10} + p_0 p_{00} = q_2,$$

$$p_1 p_{11} p_{111} + p_1 p_{10} p_{101} + p_0 p_{01} p_{011} + p_0 p_{00} p_{001} = p_3,$$
$$p_1 p_{11} p_{110} + p_1 p_{10} p_{100} + p_0 p_{01} p_{010} + p_0 p_{00} p_{000} = q_3,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

On a, de plus, d'après nos hypothèses, α, β, γ, ... désignant o ou 1,

$$\begin{aligned} p'_2 &\leqq p_{\alpha 1} \leqq p''_2, \\ p'_3 &\leqq p_{\alpha\beta} \leqq p''_3, \\ p'_4 &\leqq p_{\alpha\beta\gamma 1} \leqq p''_4, \\ &\cdots\cdots\cdots \end{aligned}$$

Les valeurs de A_0, A_1, A_2, ... sont les suivantes :

$$\begin{aligned} A_0 &= p_0 p_{00} p_{000}, \\ A_1 &= p_1 p_{10} p_{100} p_{1000} + \ldots + \ldots \\ &\quad + p_0 p_{01} h_{010} p_{0100} + \ldots + \ldots + p_0 p_{00} p_{001} p_{0010} + \ldots, \\ A_2 &= p_1 p_{11} p_{110} p_{1100} p_{11000} + \ldots \\ &\quad + p_1 p_{10} p_{101} p_{1010} p_{10100} + \ldots + p_1 p_{10} p_{100} p_{1001} p_{10010} + \ldots + \ldots \\ &\quad + p_0 p_{01} p_{011} p_{0110} p_{01100} + \ldots + \ldots \\ &\quad + p_0 p_{00} p_{001} p_{0011} p_{00110} + \ldots + \ldots . \end{aligned}$$

On voit que si l'on désigne par Π_4, Π'_4, ... les produits tels que le suivant, α, β, γ désignant o ou 1,

$$\Pi_4 = p_{\alpha\beta\gamma 0} p_{\alpha\beta\gamma 00} p_{\alpha\beta\gamma 000} \cdots,$$

on peut écrire tous les $p_{\alpha\beta\gamma}$, ... étant positifs, et par suite tous les termes des sommes A_1, A_2, ... étant positifs,

$$\begin{aligned} A_0 &= p_0 p_{00} p_{000} \Pi_4, \\ A_1 &> p_1 p_{10} p_{100} \Pi'_4 + p_0 p_{01} p_{010} \Pi''_4 + p_0 p_{00} p_{001} \Pi'''_4, \\ A_2 &> p_1 p_{11} p_{110} \Pi_4^{(4)} + p_1 p_{10} p_{101} \Pi_4^{(5)} + p_0 p_{01} p_{011} \Pi_4^{(6)}, \\ A_3 &> p_1 p_{11} p_{111} \Pi_4^{(7)}. \end{aligned}$$

Si donc on désigne par P_4 la plus petite des quantités Π_4, Π'_4, ..., $\Pi_4^{(7)}$, on a

$$A_0 + A_1 + A_2 + A_3 \geqslant (p_0 p_{00} p_{000} + \ldots + p_1 p_{11} p_{111}) P_4.$$

Mais on a évidemment

$$p_0 p_{00} p_{000} + \ldots + p_1 p_{11} p_{111} = p_3 + q_3 = 1$$

et, par suite, nous avons

$$A_0 + A_1 + A_2 + A_3 > P_4.$$

On démontrerait de même la relation

$$A_0 + A_1 + \ldots + A_k > P_{k+1},$$

P_{k+1} désignant le plus petit des produits Π_{k+1} (au nombre de 2^k) :

$$\Pi_{k+1} = p_{\alpha_1 \alpha_2 \ldots \alpha_k 0} p_{\alpha_1 \alpha_2 \ldots \alpha_k 00} p_{\alpha_1 \alpha_2 \ldots \alpha_k 000} \cdots;$$

les indices α_1, α_2, ..., α_k désignant o ou 1.

Mais on a, par exemple,

$$p_{\alpha_1\alpha_2\ldots\alpha_k 00} = 1 - p_{\alpha_1\alpha_2\ldots\alpha_k 01} \geqq 1 - p''_{k+2}$$

et, par suite,

$$\Pi_{k+1} > \prod_{n=k+1}^{\infty} (1 - p''_n) > 1 - \varepsilon.$$

Il en résulte successivement

$$P_{k+1} > 1 - \varepsilon,$$

$$A_0 + A_1 + \ldots + A_k > 1 - \varepsilon. \qquad \text{C. Q. F. D.}$$

Considérons la fraction continue, comprise entre o et 1 :

$$\cfrac{1}{a_1 + \cfrac{1}{a_2 + \ldots}},$$

et cherchons la probabilité pour qu'on ait

(2) $$a_n \geqq \varphi(n).$$

On peut se poser, à ce sujet, plusieurs problèmes.

Problème I. — *Déterminer la probabilité* A_0 *pour que l'inégalité* (2) *ne soit jamais vérifiée.*

Problème II. — *Déterminer la probabilité* A_k *pour que l'inégalité* (2) *soit vérifiée pour k valeurs de n et k seulement.*

Problème III. — *Déterminer la probabilité* A_∞ *pour que l'inégalité* (2) *soit vérifiée pour une infinité de valeurs de n.*

Les problèmes I, II et III correspondent respectivement à ceux que j'ai désignés ainsi, pages 183-186; dans le cas que j'ai appelé *cas de convergence*, les probabilités A_0, A_k ont des valeurs déterminées, ni nulles, ni infinies, et l'on a

(3) $$A_\infty = 0.$$

C'est le résultat que conteste M. Bernstein; en réalité, la probabilité qui est différente de zéro, c'est la probabilité A_0 et la probabilité complémentaire $1 - A_0$.

On peut déduire ces résultats des calculs mêmes de M. Bernstein (Be, p. 429); la formule (61) montre en effet que si n est pris assez grand, les produits infinis qui y figurent diffèrent aussi peu qu'on veut de l'unité et il en est de même de la mesure de l'ensemble des points pour lesquels l'inégalité (2) n'est vérifiée que pour un nombre fini de valeurs de n, c'est-à-dire cesse d'être vérifiée à partir d'une valeur n suffisamment grande (non fixée d'avance).

Si l'on se place au point de vue que j'ai adopté dans mon Mémoire, tout revient à démontrer que si l'on appelle *cas favorable* le cas où l'inégalité (2) est vérifiée, et si l'on suppose la série

$$\sum \frac{1}{\varphi(n)} \tag{4}$$

convergente, on se trouve dans le cas que j'ai appelé *cas de convergence*.

Or, la probabilité $Q_{n,k}$, pour qu'on ait

$$a_n \geqq k$$

$\left(\text{probabilité que j'ai appelée } 1 - P_{n,k} \text{ et que M. Bernstein appelle } \frac{C_{n,k}}{C_{n,1}}\right)$, vérifie des inégalités de la forme

$$\frac{A}{k} < Q_{n,k} < \frac{B}{k}, \tag{5}$$

A et B étant des constantes indépendantes de k [Be, formule (42), p. 426; *suprà*, formules (23) et suivantes, p. 205]. M. Bernstein fait observer très justement que cette probabilité $Q_{n,k}$ n'est pas indépendante des valeurs de $a_1, a_2, \ldots, a_{n-1}$; mais, *quelles que soient ces valeurs*, elle vérifie les inégalités (5); comme toutes les démonstrations, celle de M. Bernstein comme la mienne, sont basées, non sur la valeur exacte et très compliquée de $Q_{n,k}$, mais sur les inégalités (5), ces démonstrations ne sont en rien modifiées par le fait que $Q_{n,k}$ est variable.

Si l'on suppose $k = \varphi(n)$, on posera

$$Q_{n,\varphi(n)} = p_n$$

et l'on trouvera la notation avec laquelle j'ai traité les problèmes I, II et III (p. 183-186); la série

$$\sum p_n$$

est bien convergente puisqu'elle se réduit à la série (4) multipliée par un facteur inconnu, mais compris entre les nombres fixes A et B. On peut dire aussi que p'_n et p''_n vérifient les inégalités (5). Le fait que les p_n ne sont pas constants entraîne comme on l'a vu une modification de forme dans la démonstration, mais n'altère pas le résultat essentiel : si la série (4) est *convergente*, la probabilité pour qu'on ait, à partir d'une certaine valeur de n,

$$a_n > \varphi(n)$$

est égale à zéro; il y a une probabilité égale à *un* pour que l'inégalité

$$a_n < \varphi(n)$$

soit vérifiée, non pour *toute* valeur de n, mais pour les valeurs de n suf-

fisamment grandes. En d'autres termes, il y a une probabilité égale à 1 pour que a_n soit *asymptotiquement inférieur* à $\varphi(n)$. Il y a de même une probabilité égale à *un* pour que a_n soit *asymptotiquement supérieur* à $\varphi(n)$, si la série (4) est *divergente*. De sorte qu'il est infiniment probable que la croissance asymptotique de a_n est comprise entre toute fonction $\varphi(n)$ *donnée* telle que la série (4) soit divergente et toute autre fonction *donnée* telle que cette série soit convergente.

Puisque l'occasion m'a été donnée de parler du Mémoire de M. Bernstein, je voudrais dire aussi quelques mots de son *Axiome* (Be, p. 419). J'ai souvent pensé à des considérations de ce genre et je suis convaincu, avec M. Bernstein, que la théorie de la mesure, et en particulier de la mesure nulle, est appelée à jouer un rôle important dans les questions de mécanique statistique. La difficulté essentielle dans ces problèmes est en effet l'existence possible de mouvements ordonnés, qui sont très peu probables, mais dont la probabilité n'est pas rigoureusement égale à zéro. En particulier, pour le théorème du minimum de H. de Boltzmann, il est bien connu que les mouvements tels que H croisse, soit dans l'*avenir*, soit dans le *passé* à partir de sa valeur actuelle, sont infiniment moins probables que les mouvements tels que H décroisse (ou conserve sa valeur minimum, si elle est atteinte). En réalité, si le nombre des molédules n'est pas *infini*, l'ensemble des trajectoires correspondant à la croissance de H n'est pas de mesure *nulle;* sa mesure est seulement extrêmement petite. Il suffira donc d'une perturbation extérieure excessivement faible, d'une action stellaire sur un système terrestre, par exemple, pour que l'irréversibilité ne soit pas possible [1]. C'est en ce sens que je comprends l'axiome de M. Bernstein ; il revient à ceci : *Pratiquement, une probabilité nulle ou extrêmement petite doit être considérée comme équivalente à l'impossibilité.* Nous entendons par *extrêmement petite* une probabilité telle que l'événement attendu doive se produire une fois, en moyenne, dans l'univers accessible à l'homme au cours d'une période de temps très grande par rapport à la durée du système solaire.

---

[1] *Voir* mon article *Sur les principes de la théorie cinétique des gaz* (*Ann. École Normale*, 1906) et ma Note : *Modèles arithmétiques et analytiques de l'irréversibilité apparente* (*C. R. Acad. Sc.*, t. CLIV, p. 1148).

NOTE VI.

LA THÉORIE DE LA MESURE ET LA THÉORIE DE L'INTÉGRATION.

INTRODUCTION.

I. — *Remarques historiques.*

Les résultats acquis, dès la fin du XIXe siècle, ont surabondamment prouvé combien était simpliste l'opinion d'après laquelle il serait possible de limiter le champ des Mathématiques à l'étude d'une catégorie déterminée de fonctions : fonctions continues, fonctions dérivables, fonctions analytiques, etc. Pour qu'une telle limitation ne fût pas à la fois arbitraire et vaine, il faudrait, en effet, qu'on pût être assuré de son invariance, à l'égard du moins d'une catégorie déterminée de transformations analytiques. Or, si l'on n'a pas le droit d'affirmer qu'une telle limitation sera toujours impossible, on doit reconnaître que sa réalisation prochaine est peu vraisemblable dans l'état actuel de la Science. Cette réalisation exigerait, entre autres choses, une étude approfondie au point de vue arithmétique de tous les nombres irrationnels qui peuvent s'introduire en Algèbre et en Analyse, et une telle étude est à peine ébauchée (¹). En effet, l'introduction d'un nombre irrationnel α d'une complication particulière dans une équation fort simple telle que la suivante

$$\frac{\partial^2 z}{\partial x^2} - \alpha \frac{\partial^2 z}{\partial y^2} = \varphi(x, y)$$

entraîne une conséquence inattendue : en partant de conditions *analytiques*, l'équation définit une fonction des variables x et y *qui n'est analytique pour aucun système de valeurs de ces variables* (²). Dans un autre ordre d'idées, M. Lebesgue a tiré, de la considération des développements décimaux des nombres irrationnels les plus généraux, des conséquences presque paradoxales; en particulier il en a déduit la définition

(¹) *Voir*, par exemple, mes *Leçons sur la Théorie de la croissance*, dernier Chapitre.

(²) BOREL, *C. R. Acad. Sc.*, t. CXXI, 19 décembre 1895.

d'une fonction qui n'est susceptible d'aucune représentation analytique ([1]).

Cette quasi-impossibilité d'établir une démarcation précise entre les êtres analytiques regardés comme « simples » et les autres, a été l'origine de travaux qui ont considérablement accru nos connaissances en Analyse. Ces travaux étaient nécessaires; ils ne sont pas d'ailleurs définitifs sur tous les points et il sera encore utile, à mon avis, de s'occuper de ce qu'on a pu appeler la *pathologie* des fonctions. Mais il est permis de penser que le but définitif de ces recherches *pathologiques* doit être la délimitation des fonctions considérées comme *saines* ([2]). Là encore, nous nous heurtons à des difficultés qui sont loin d'être résolues.

Je voudrais essayer d'indiquer un point de vue différent, qui n'est pas entièrement nouveau, mais qu'il m'a semblé possible, en utilisant divers travaux publiés dans ces vingt dernières années, de présenter sous une forme qui me paraît neuve. J'en donnerai, d'ailleurs, immédiatement une application concrète au calcul des intégrales définies, de manière à bien montrer qu'il ne s'agit pas d'une pure théorie spéculative, mais que la conception que je propose peut conduire à des résultats positifs et précis, indépendamment de toute opinion sur cette conception même.

L'idée qui m'a guidé est l'utilité qui me paraît évidente de distinguer entre les calculs qui peuvent être réellement effectués et ceux qui ne peuvent pas l'être. Les premiers, seuls, sont actuellement utilisables dans les applications des Mathématiques. Je ne veux pas dire, bien entendu, que les applications soient l'unique but des Mathématiques; rien n'est plus loin de ma pensée; et, admettrait-on ce point, il n'en resterait pas moins que certaines spéculations, aujourd'hui sans rapport visible avec aucune application, se révéleront peut-être demain comme très fécondes en résultats pratiques. Ce que je dis simplement, c'est qu'il y a un grand intérêt théorique et pratique à étudier à part les nombres et les fonctions *calculables;* nous verrons que ce champ d'étude est beaucoup plus étendu qu'on n'aurait pu le penser il y a quelques années; c'est pourquoi je me suis décidé à cette publication, à laquelle je réfléchis depuis fort longtemps.

Ces réflexions, je n'ai pas besoin de le dire, n'ont pu être indépendantes de mes lectures, ni surtout de mes conversations. Il ne m'est pas possible de signaler tous ceux qui, par leurs paroles ou par leurs écrits, ont eu une part dans la formation des idées que je vais exposer : je ne les connais d'ailleurs pas tous, car ces influences sont parfois inconscientes. Je manquerais cependant à un devoir élémentaire de probité, si je ne signa-

---

([1]) *Sur les fonctions représentables analytiquement* (*Journal de M. Jordan*, 1905, p. 214). Sur le sens qu'il faut donner au mot *définition*, M. Lebesgue fait des observations fort justes qui entraînent des réserves sur le sens que peut avoir son énoncé, dont l'intérêt, à mon point de vue, est surtout négatif.

([2]) J'ai déjà exprimé, à plusieurs reprises, cette opinion, en particulier en ce qui concerne la régularité des modes de croissance (*Leçons sur les fonctions entières*, Note III).

lais pas certaines conversations amicales dont il n'est pas resté de trace écrite et qui ont certainement joué un rôle essentiel : je veux parler de conversations avec M. Jules Drach qui remontent à plus de vingt ans et ont été souvent reprises depuis; et ensuite, dans l'ordre chronologique, de conversations avec M. René Baire et avec M Henri Lebesgue. Je conserve, cela va sans dire, toute la responsabilité de ce qui pourra paraître critiquable dans les considérations auxquelles je consacre cette Introduction; mais je voudrais, si quelqu'une d'entre elles retient l'attention, qu'on sache qu'elle n'aurait probablement pas vu le jour sous sa forme actuelle si je n'avais pas eu la bonne fortune d'échanger des idées avec les amis dont je viens de donner les noms.

II. — *Nombres calculables.*

Nous dirons qu'un nombre α est calculable lorsque, étant donné un nombre entier quelconque n, on sait obtenir [1] un nombre rationnel qui diffère de α de moins de $\frac{1}{n}$. Les nombres rationnels sont évidemment les plus simples des nombres calculables; lorsqu'un nombre n'est par rationnel, on en calcule généralement les valeurs approchées, en faisant usage, soit des fractions décimales, soit des fractions continues. On peut se demander s'il faut considérer à part ceux des nombres irrationnels pour lesquels *la loi* d'un de ces développements est connue; il est clair qu'il en résulte un avantage pratique considérable, mais il faut observer que cet avantage est entièrement limité à un mode unique de représentation ; le système décimal, en particulier, n'a aucune valeur théorique spéciale; il n'en est pas de même pour le développement en fraction continue, qui est unique en son genre, mais, d'autre part, un tel développement n'est pas invariant relativement à des opérations arithmétiques très simples [2].

---

[1] Je laisse intentionnellement de côté la plus ou moins grande longueur pratique des opérations; l'essentiel est que chacune de ces opérations soit exécutable en un temps fini, par une méthode sûre et sans ambiguïté. Par exemple un nombre décimal β, tel que sa $n^{\text{ième}}$ décimale soit égale à la décimale de π de rang $n!$ doit être regardé comme défini, bien que son calcul, avec seulement une dizaine de chiffres exacts, puisse exiger, dans l'état actuel de l'Analyse, un temps dépassant de beaucoup la longueur de la vie humaine. En réalité, la difficulté est la même pour tous les nombres incommensurables; les premières décimales sont parfois aisées à calculer, mais l'impossibilité pratique reparaît si l'on exige quelques milliers de chiffres. Au point de vue pratique, on peut dire que les nombres dont on a effectivement besoin peuvent, en général, être effectivement calculés avec l'approximation désirable; d'autre part, il n'y a pas lieu d'élever des exigences de nature pratique, lorsqu'il s'agit de nombres dont l'importance pratique est nulle, comme le nombre β dont il vient d'être question.

[2] Quelques recherches relatives à cette invariance ont été entreprises par M. Châtelet (*Bulletin de la Société mathématique*, 1912, t. XL, p. 1). Mais, malgré l'intérêt des résultats obtenus, ceux-ci sont encore très particuliers et

Il en résulte que la connaissance de la loi d'un développement décimal (non périodique) doit être actuellement considérée comme n'ayant aucune valeur théorique [1], ni pratique, tandis que la connaissance de la loi d'un développement en fraction continue a un certain intérêt theorique, mais un intérêt pratique à peu près nul.

Le premier des problèmes qui se pose au sujet des nombres calculables est celui de l'égalité de deux de ces nombres [2]. Si deux nombres calculables sont inégaux, on s'en apercevra évidemment en calculant chacun d'eux avec une approximation suffisante, déterminée, mais généralement inconnue *a priori*. On réaliserait un progrès évident en déterminant une limite inférieure de la différence qui peut exister entre deux nombres calculables, dont les définitions satisfont à des conditions connues. Cette limite existe évidemment si les conditions sont telles qu'elle ne définissent qu'un nombre fini de nombres calculables; j'ai déjà indiqué ailleurs cette manière de poser la question [3]; je veux seulement faire observer ici que la fonction qui définit l'écart minimum des deux nombres de *hauteur* donnée [4] est calculable, pour chaque valeur finie de la hauteur, mais que son calcul effectif serait d'une longueur impraticable, s'il fallait le réaliser empiriquement, c'est-à-dire en calculant tous les nombres dont la hauteur ne dépasse pas la valeur finie considérée. On n'est même pas absolument sûr, au point de vue théorique, qu'il ne se présenterait pas des difficultés insolubles, car on peut concevoir deux nombres tels que les suivants :

$$\left(\int_{-\infty}^{+\infty} e^{-x^2}\,dx\right)^2, \quad \int_{-\infty}^{+\infty} \frac{dx}{1+x^2},$$

---

leur extension paraît présenter de grandes difficultés. C'est là une des questions les plus importantes de l'Arithmétique, et il serait très désirable que les recherches de M. Châtelet soient continuées et étendues.

(1) Je suppose ici, bien entendu, que cette loi serait *tout* ce qu'on saurait sur le nombre. Au contraire, il y aurait un très grand intérêt théorique à connaître la loi des chiffres décimaux d'un nombre défini autrement que par cette loi, de $\sqrt{2}$ par exemple. Mais c'est là un sujet de recherches très difficile, et qui ne pourra être abordé qu'après celui dont il vient d'être question dans la note précédente.

(2) La connaissance des lois dont il vient d'être question permettrait de résoudre cette question pour deux nombres donnés *sous la même forme*, mais le problème n'a alors aucun intérêt. Ce qui importe, c'est de savoir reconnaître si deux nombres sont égaux, lorsqu'ils sont obtenus par des modes de calcul différents.

(3) *C. R. Acad. Sc.*, décembre 1903.

(4) Je rappelle que la hauteur d'un nombre est une fonction croissante du nombre d'opérations nécessaires pour définir le nombre à partir de l'unité; la définition est seulement assujettie aux conditions suivantes : 1° tout nombre calculable doit avoir une hauteur finie; 2° il y a un nombre fini de nombres dont la hauteur est inférieure à une valeur donnée.

ayant un nombre quelconque de chiffres décimaux identiques et dont on ne saurait pas prouver l'égalité (1). Il serait donc tout à fait désirable qu'on ait quelques résultats généraux non empiriques sur la limitation inférieure de l'écart en fonction de la hauteur (2). C'est un problème qui mérite d'attirer l'attention des analystes, malgré sa difficulté.

Un autre problème, bien plus difficile que le précédent et cependant plus souvent signalé comme digne d'intérêt, consiste à se demander si un nombre calculable appartient ou non à une catégorie énumérable (3) particulière : si le nombre π, ou la constante d'Euler C est ou non un nombre rationnel, ou quadratique, ou algébrique? On sait que la question est résolue pour π, et ne l'est pas pour C. Il ne semble pas possible, actuellement, d'aborder ce problème sous sa forme générale, pour les nombres qui peuvent être définis, à partir des nombres rationnels, par des conditions transcendantes. On peut le ramener au problème, à certains égards plus simple, qui consiste à *limiter la hauteur* des nombres rationnels qui peuvent être définis à partir des nombres entiers par des opérations algébriques ou analytiques de nature déterminée; il est clair que si les procédés de définition sont tels qu'ils ne conduisent qu'à un nombre *fini* de nombres, la hauteur de ceux de ces nombres qui sont rationnels est limitée : mais cette remarque bien simple, si elle permet d'entrevoir la possibilité d'une solution, ne fournit, malheureusement, aucune méthode pour l'obtenir.

Je n'insiste pas davantage sur ces difficultés qui sont aux frontières du sujet qui nous occupe; je ne ferai que mentionner d'autres difficultés qui, à mon sens du moins, sont au delà des frontières des Mathématiques. Je fais allusion aux *définitions* telles que la suivante, qu'on peut varier à l'infini : le nombre a est égal à *zéro* si la constante d'Euler C est un

---

(1) Il ne semble pas que ces difficultés se présentent *effectivement* dans la réalité; pratiquement, toutes les fois que les mathématiciens ont constaté l'égalité *numérique* de deux nombres (à une approximation suffisante), ils ont su démontrer l'égalité rigoureuse. C'est là un fait important : dans tous les cas où l'on sait définir deux nombres dont les développements décimaux ont une centaine de chiffres communs, on sait aussi, ou bien qu'ils sont égaux, ou bien qu'ils sont inégaux en vertu de leur définition même (par exemple, l'un d'eux peut être défini comme égal à l'autre augmenté de 10^{-1000}). Il serait fort intéressant de pouvoir donner un exemple de deux nombres dont les développements décimaux coïncident *pratiquement*, sans qu'on sache s'ils sont égaux on non. Ceci se rattache à la question de savoir dans quelle mesure une vérification empirique assure de l'exactitude d'un théorème d'Analyse, tel que le fameux théorème de Riemann sur la fonction $\zeta(s)$.

(2) *Voir* mes *Leçons sur la théorie de la croissance* (*loc. cit.*).

(3) J'emploie le terme *énumérable* dans le sens que j'ai donné aux mots : *effectivement énumérable*, réservant au terme *dénombrable* son sens usuel, sans m'inquiéter ici des critiques qu'on peut faire à cette notion d'ensemble dénombrable non effectivement énumérable (voir *Annales de l'École Normale*, 1908 : *Les paradoxes de la théorie des ensembles*).

nombre algébrique et à *un* dans le cas contraire. En d'autres termes, on fait dépendre la valeur du nombre *défini* d'une certaine éventualité inconnue; la seule raison pour laquelle on regarde la *définition* comme mathématique est que l'éventualité inconnue est de nature mathématique et que, par suite, le nombre a est susceptible d'une *définition* analytique; il suffirait d'un peu de patience pour écrire explicitement une formule donnant ce nombre a; mais cette formule renfermerait plusieurs passages à la limite superposée et ne serait évidemment pas calculable. La *définition* analytique n'a donc aucune valeur mathématique; elle est simplement la traduction, en un langage plus compliqué, de la définition primitive, de sorte qu'on est, de toute façon, ramené à faire dépendre la valeur du nombre a de la solution d'un problème pour le traitement duquel on ne possède aucune méthode régulière. Le fait que ce problème est mathématique me paraît être une circonstance accessoire et il me semble qu'on aurait une définition tout à fait analogue en disant que le nombre a est égal à *zéro* ou à *un*, suivant que le cuivre est ou non un corps composé. Il est, en effet, aussi peu vraisemblable pour les mathématiciens que C soit algébrique, qu'il est peu vraisemblable pour les chimistes que le cuivre soit un corps composé; mais la preuve rigoureuse paraît presque également difficile dans les deux cas [1].

Nous nous en tiendrons donc à la définition des nombres calculables donnée au début de ce paragraphe: les commentaires dont nous l'avons fait suivre n'avaient comme but que de mettre en évidence les restrictions que comporte une telle définition; mais ces restrictions n'ont rien d'arbitraire; elles s'imposent si l'on veut distinguer les mathématiques réelles de spéculations *logiques* purement verbales, dans lesquelles on ne se préoccupe que d'une qualité purement négative : l'absence de contradiction.

III. — *Les fonctions calculables et les fonctions à définition asymptotique.*

Nous dirons qu'une fonction est calculable lorsque sa valeur est calcu-

---

[1] On pourrait objecter à cet exemple, ainsi qu'à tout autre exemple tiré des sciences physiques, les difficultés possibles d'interprétation : la notion de corps simple peut être entièrement modifiée; la continuité des phénomènes est une autre difficulté qui paraît être très profonde : entre deux alternatives naturelles, il y a toujours place pour le doute : le seul moyen sûr d'échapper au doute est de substituer au phénomène naturel un phénomène au moins en partie conventionnel et artificiel. Par exemple, on ne peut pas parler avec certitude du nombre des petites planètes qui seront découvertes avant le 31 décembre 1920, à minuit (temps de Paris), car une découverte peut avoir lieu précisément à minuit; et, d'autre part, il peut y avoir doute sur la valeur d'une observation particulière; mais on peut parler avec précision des découvertes qui auront été annoncées à cette date dans une publication déterminée, si cette publication, par suite de conventions humaines, paraît à des dates régulières telles qu'il ne puisse pas y avoir d'ambiguïté à redouter.

lable pour toute valeur calculable de la variable [1]. En d'autres termes, si α est un nombre calculable, on doit savoir calculer la valeur de $f(\alpha)$ à $\frac{1}{n}$ près, quel que soit n. On ne doit pas perdre de vue que, par définition, donner le nombre calculable α, c'est simplement donner le moyen d'obtenir α avec une approximation arbitraire. Une fonction ne peut donc être calculable que si elle est continue [2], au moins pour les valeurs calculables de la variable.

Si l'on donne les valeurs d'une fonction pour les valeurs calculables de la variable et si cette fonction est supposée continue, ses valeurs pour *toutes* les valeurs de la variable sont par cela même *déterminées*. On peut se demander si l'on peut attribuer un sens quelconque à la valeur d'une fonction complètement discontinue pour les valeurs non calculables de la variable, même si cette fonction était continue dans le champ des valeurs calculables. Il semble qu'on doive chercher à écarter, *a priori*, une singularité artificielle analogue à la suivante : une fonction serait égale à x pour x calculable et à x^2 pour x non calculable. En d'autres termes, la valeur de la fonction pour les valeurs non calculables serait égale à la limite d'une certaine fonction continue bien définie pour les valeurs calculables. Une telle convention apparaît, en effet, comme une discontinuité artificielle, analogue à celle qui consisterait à considérer une fonction de la variable complexe z, égale dans tout le plan à z, sauf pour $z = 0$, où sa valeur serait 27; on convient généralement de laisser de côté de telles fonctions qui ne posséderaient pas la propriété fondamentale des fonctions analytiques : deux fonctions analytiques qui coïncident dans le voisinage d'un point ont même domaine d'existence et coïncident dans tout ce domaine. De même, il peut sembler naturel de convenir qu'une fonction continue pour toutes les valeurs calculables doit être considérée comme continue pour toutes les valeurs non calculables; elle est, par cela même, *définie* pour ces valeurs autant qu'elle peut l'être.

Aux fonctions calculables, on peut opposer *les fonctions à définition asymptotique*. Je propose d'appeler ainsi les fonctions dont la valeur, pour une valeur déterminée de la variable, ne dépend que de la manière dont se comporte *à l'infini* un développement convergent de cette valeur de la variable. C'est là le type qui est le plus éloigné des fonctions calculables; on pourrait évidemment concevoir des types mixtes; je ne m'y attarderai pas.

Un nombre peut être considéré comme défini par une suite énumérable

---

(1) Nous ne parlons que d'une variable pour abréger le langage; il ne se présente aucune difficulté pour l'extension à n variables ou même à une infinité énumérable, sous certaines restrictions de convergence évidentes.

(2) Pour que le calcul de la fonction soit effectivement possible à une approximation donnée, il faut, de plus, supposer connue la mesure de la continuité de la fonction, c'est-à-dire l'ordre infinitésimal (au sens généralisé) de la variation de la fonction comparée à la variation de la variable.

d'entiers, tels que deux nombres soient infiniment voisins si, pour des valeurs de plus en plus grandes de n, les n premiers entiers de la suite coïncident pour les deux nombres. On peut supposer que la valeur de la fonction est en partie déterminée par les n premiers entiers, ou, au contraire, que, quel que soit n, elle ne dépend pas de ces n premiers entiers; ce sont les deux types extrêmes de la fonction continue et de la fonction à définition purement asymptotique. Pour compliquer encore, on pourrait essayer de ranger les entiers sous la forme d'une suite bien ordonnée, correspondant à un nombre transfini de deuxième classe (variable) et admettre que, quel que soit le nombre α de deuxième classe, il est des valeurs de x telles que $f(x)$ ne dépende pas de la valeur des α premiers termes de la suite. Mais bien des réserves seraient à faire sur la légitimité d'une telle définition; je me bornerai aux définitions asymptotiques simples.

La plus classique des fonctions à définition asymptotique est la fonction souvent considérée qui est égale à o ou à 1, suivant que la variable x est un nombre rationnel ou irrationnel; nous y reviendrons tout à l'heure. On peut déduire des développements des nombres en fraction décimale ou en fraction continue bien des définitions plus ou moins compliquées, dans lesquelles intervient la loi asymptotique de ces développements. Un tel développement définit une suite énumérable de nombres entiers; on peut considérer, par exemple, une fonction de plusieurs nombres de cette suite dont les rangs varient suivant une certaine loi et envisager la plus grande ou la plus petite limite vers laquelle tend cette fonction, lorsque les rangs des nombres qui y figurent augmentent indéfiniment. C'est là un type déjà très général et qu'on pourrait encore compliquer. Je n'ai pas l'intention d'aborder ici l'étude générale de telles fonctions; je voudrais seulement faire observer que cette étude me paraît être du ressort de la théorie des probabilités. En effet, on ne peut se borner aux valeurs calculables de la variable; et une valeur non calculable ne peut être concue comme définie que par le hasard; les propriétés de la fonction sont donc représentées par des coefficients de probabilité.

Un cas particulier important est celui où la fonction coïncide avec une fonction continue pour toutes les valeurs non calculables et diffère de cette fonction seulement pour certaines valeurs calculables. Tel est le cas d'une fonction égale, pour x irrationnel, à 1 ou à x et, pour x rationnel, à zéro ou à une fonction déterminée du numérateur et du dénominateur de la fraction irréductible égale à x. Dans ce cas, il est clair que, si un nombre est donné par une série dont on ignore la loi, il y a une probabilité égale à l'unité pour que ce nombre soit irrationnel, et l'on doit, par conséquent, choisir, dans le doute, comme valeur approchée de la fonction, la valeur qui correspond à cette hypothèse, infiniment plus probable que l'hypothèse contraire.

Une fonction à définition asymptotique est parfois immédiatement connue pour un grand nombre de valeurs calculables, lorsque ces valeurs sont données sous une forme particulière qui peut être, suivant les cas, arithmé-

tique, algébrique ou analytique. Par exemple, on peut convenir qu'une fonction de deux variables x et y est nulle, sauf si $x+iy$ est égale à une période d'une fonction $p(u, g_2, g_3)$ à invariants rationnels, auquel cas la fonction est égale à 1. Il est clair que, si l'on se donne des nombres rationnels g_2 et g_3, ils définissent certains systèmes de valeurs x et y pour lesquels la fonction est connue. Mais si l'on donne, par une autre voie, un système de nombres x et y, nous ne connaissons actuellement aucun moyen de déterminer la valeur correspondante de la fonction [1]; le problème ainsi posé n'a qu'un rapport assez lointain avec la définition de la fonction; il mérite d'être étudié en lui-même, indépendamment de cette définition. On pourrait concevoir de même des fonctions non calculables à propos desquelles on serait ainsi amené à se poser les problèmes les plus divers; en ce sens, la théorie des fonctions non calculables comprendrait la Science tout entière [2] : c'est une raison sans doute suffisante pour ne pas l'aborder dans sa généralité; nous verrons plus loin sous quelles conditions certaines fonctions non calculables peuvent être maniées.

IV. — *Les ensembles mesurables.*

La notion d'ensemble est un cas particulier de la notion generale de fonction; tout ensemble définit une fonction, égale à zéro pour les points d'ensemble et égale à 1 pour les points qui n'appartiennent pas à l'ensemble. Inversement, toute fonction égale à *zéro* ou à *un* pour les points d'un domaine sépare les points de ce domaine en deux catégories, c'est-à-dire définit deux ensembles complémentaires. Une fonction quelconque définit une courbe, c'est-à-dire un ensemble particulier.

Si une fonction ne prenant que les valeurs 0 et 1 est continue en un point, elle est évidemment constante dans un intervalle contenant ce point; l'étude des fonctions continues conduit donc à celle des ensembles formés d'intervalles, les extrémités de ces intervalles jouant le rôle des points de discontinuité. On peut, d'ailleurs, considérer soit un ensemble E formé d'intervalles, soit l'ensemble E' complémentaire de E, soit enfin combiner les deux modes de formation; nous reviendrons sur ce point.

Je voudrais auparavant faire observer que, si l'on applique à la définition des ensembles les remarques faites sur la définition des fonctions et qu'on appelle *bien définis* les ensembles qui correspondent aux fonctions calculables, on voit que les seuls ensembles *bien définis* sont ceux dont la définition peut être obtenue au moyen d'intervalles (additifs ou soustractifs), les *extrémités* des intervalles devant être étudiées à part. Si l'on ne tient

(1) Nous savons simplement calculer g_2 et g_3 avec une approximation indéfinie; c'est une propriété asymptotique de ces nombres d'être rationnels ou non.

(2) On peut dire en effet : telle fonction est égale à 0 ou à 1 pour $x=0$, suivant que telle proposition (par exemple le dernier théorème de Fermat) est vraie ou fausse. Il serait facile de traduire cet énoncé par une formule.

pas compte de ces extrémités, on peut dire que l'ensemble est défini à un ensemble dénombrable près ([1]). Relativement aux ensembles énumérables, les difficultés soulevées par leur définition sont différentes suivant qu'ils sont réductibles ou denses dans un intervalle. Le cas le plus simple de l'ensemble réductible est l'ensemble comprenant un seul point a. Si un point x quelconque est donné par un procédé quelconque, et si x diffère de a, on s'en apercevra au bout d'un nombre d'opérations fini (mais non connu d'avance); si x coïncide avec a, la démonstration de ce fait ne résultera pas en général d'un calcul d'approximation fini, mais des définitions théoriques de x et de a (*Cf.* p. 220.) L'ensemble formé d'un seul point, serait-ce le point *zéro*, n'est donc pas bien défini en ce sens que le problème de savoir si un nombre donné appartient ou non à l'ensemble peut exiger, soit une infinité d'opérations, soit la résolution d'un problème difficile ou actuellement insoluble. Les difficultés sont plus grandes encore lorsqu'il s'agit d'un ensemble dense, par exemple de l'ensemble des nombres rationnels; étant donné un nombre x défini analytiquement, on ne sait généralement pas reconnaître s'il appartient ou non à l'ensemble ([2]).

Si l'on néglige les ensembles dénombrables, les ensembles *bien définis* sont précisément les ensembles que j'ai appelés autrefois *ensembles mesurables* ([3]). Depuis, M. Lebesgue a donné à ces ensembles le nom d'*ensembles mesurables* B et a considéré d'autres ensembles qu'il a nommés *ensembles mesurables*. J'adopterai cette dernière dénomination de M. Lebesgue, réservant le nom d'ensembles *bien définis*, pour les ensembles que j'avais appelés *mesurables*. La terminologie que j'avais adoptée dans mes *Leçons sur la théorie des fonctions* aurait pu, en effet, avoir l'inconvénient de laisser croire qu'on ne pouvait pas parler de la mesure d'un ensemble ne rentrant pas dans la catégorie des ensembles que j'appelais alors *mesurables* et que j'appelle maintenant *bien définis;* j'ai cependant indiqué, dans ce Livre même, en vertu de quels principes la définition de la mesure pouvait s'étendre à certains ensembles qui n'étaient pas bien définis [c'est-à-dire non mesurables, avec le langage que j'adoptais alors ([4])]; et j'ai indiqué une application de ces principes à un ensemble

---

([1]) Cet ensemble est même *effectivement énumérable*, car il est toujours possible, en tenant compte des longueurs et des situations respectives des intervalles, de fixer un ordre précis dans lequel on les supposera rangés.

([2]) En réalité, la difficulté est *la même* pour *tout* ensemble dénombrable, qu'il se compose d'un seul point ou d'une infinité, lorsqu'on envisage le problème dans toute sa généralité, car il est aisé de former une expression analytique y qui s'annule dans le cas où x appartient à l'ensemble dénombrable et dans ce cas seulement. La question de savoir si x appartient à cet ensemble équivaut donc à celle de savoir si y appartient à l'ensemble formé du seul point *zéro*. Mais x peut être calculable et y non calculable; c'est la différence entre les deux cas.

([3]) *Leçons sur la théorie des fonctions*, Chap III.

([4]) « Si un ensemble E contient tous les éléments d'un ensemble mesurable E_1 de mesure α, nous pouvons dire que la mesure de E est supérieure à α, sans nous

de mesure nulle (¹). Lorsque M. Lebesgue a donné sa définition, il a eu soin de faire observer que tout ensemble mesurable E était compris dans un ensemble bien défini E_1 (*mesurable* B, dans le langage de M. Lebesgue) et comprenait un ensemble bien défini E_2, les ensembles E_1 et E_2 ayant même mesure. Ceci entraîne que la mesure de E est égale à la mesure commune de E_1 et de E_2; la mesure des ensembles mesurables se déduit donc de celle des ensembles bien définis, et la classe des ensembles, dont la mesure peut être déduite de mes définitions primitives, est identique à la classe des ensembles, dont la mesure peut être déduite des définitions de M. Lebesgue. Ce qui différencie la définition de M. Lebesgue de la mienne, ce n'est donc pas l'étendue de la catégorie des ensembles auxquels ces définitions s'appliquent (*voir* ci-dessous, Chap. I, § V), c'est le fait que M. Lebesgue indique une marche théorique pour obtenir la mesure d'un ensemble mesurable défini d'une manière quelconque; tandis que je m'étais occupé seulement d'arriver à la mesure de certains ensembles, en utilisant la définition de ces ensembles qui se présentait naturellement à l'occasion de mes recherches de théorie des fonctions. Je pense que si l'on se place au point de vue que j'ai exposé dans cette Introduction, la plus grande généralité ainsi obtenue par M. Lebesgue est plus apparente que réelle; la théorie de la mesure des ensembles, et aussi celle du calcul des intégrales définies, me paraît être exposée d'une manière plus simple et plus élémentaire et, en même temps, aussi réellement générale, en suivant la marche que j'avais primitivement adoptée. C'est le mode d'exposition que je suivrai; il me conduirait logiquement à passer sous silence les travaux de M. Lebesgue. Mais aucun lecteur ne s'y trompera; sans qu'il soit nécessaire que je redise ici toute l'estime scientifique que j'ai pour M. Lebesgue (²), il est évident, pour tous, que ses idées ont réagi sur les miennes et que, même si la définition de l'intégrale que je propose paraît plus simple et plus naturelle, c'est à lui que la Science restera redevable de quelques-uns des progrès les plus essentiels que la théorie des fonctions ait faits dans ces dix dernières années.

---

inquiéter si E est mesurable ou non. Inversement, si E_1 contient tous les éléments de E, nous dirons que la mesure de E est inférieure à α. Les mots *supérieure* et *inférieure* n'excluent d'ailleurs pas l'égalité » (*Op. cit.*, p. 48).

(¹) « L'ensemble des points de divergence a pour mesure zéro, nous ne sommes pas assurés que cet ensemble soit mesurable; en employant le langage de la page 48, nous devrions dire que la mesure est inférieure ou égale à zéro; mais la mesure n'est jamais négative » (*Op. cit.*, p. 67).

(²) *Voir* mon article de la *Revue générale des Sciences*, t. XX, 1909, p. 315 *La théorie des ensembles et les progrès récents de la théorie des fonctions.*

CHAPITRE I.

La théorie de la mesure.

J'adopterai un mode d'exposition synthétique, en ne supposant connues que les parties classiques élémentaires de l'Analyse [1]; je serai ainsi amené à revenir sur certaines propositions que j'ai déjà démontrées ou énoncées; mais j'en donnerai le plus souvent des démonstrations nouvelles et plus simples.

I. — *Le premier théorème fondamental.*

La théorie de la mesure est basée sur le théorème suivant, que nous appelons premier théorème fondamental [2].

Si l'on a, sur un segment de droite, une infinité dénombrable d'intervalles, tels que tout point de la droite soit intérieur à au moins l'un d'eux, on peut choisir parmi eux un nombre limité d'intervalles ayant la même propriété. Soit ab le segment donné; tout point x de ab est, par hypothèse, *intérieur* à un intervalle $a_n b_n$, c'est-à-dire est compris entre a_n et b_n sans coïncider avec ces points; le point a coïncide avec l'extrémité a_i d'un, au moins, des segments $a_i b_i$ et le point b avec l'extrémité b_j d'un, au moins, des segments $a_j b_j$ (on suppose que les points a, a_n sont respectivement à gauche des points b, b_n).

Cela étant, choisissons parmi les intervalles $a_i b_i$, tels que a_i coïncide avec a, celui pour lequel l'indice i est le moins élevé; soit a_{n_1}, b_{n_1}; soit, de même, a_{n_2} b_{n_2} l'intervalle d'indice minimum renfermant le point b_{n_1}, et généralement a_{n_k} b_{n_k} l'intervalle d'indice minimum renfermant $b_{n_{k-1}}$. Je dis qu'on atteint le point b au bout d'un nombre fini d'opérations, c'est-à-dire qu'il existe un nombre fini k tel que b_{n_k} coïncide avec b. Car si les b_{n_k} étaient un nombre infini, ils admettraient un point limite b' (qui pourrait être b); à ce point b', correspond un intervalle $a_n b_n$, tel que a_n soit à gauche de b' et b_n à droite de b' (ou coïncidant avec b' si b' est b); les b_{n_k} seraient donc compris à partir d'une certaine valeur de k à l'intérieur de l'intervalle

---

(1) *Voir* la Préface de mes *Leçons sur la théorie des fonctions.*

(2) On a donné parfois à ce théorème le nom de théorème de Heine-Borel, en raison de son analogie avec la démonstration donnée par Heine du fait que la continuité est uniforme (*Journal de Crelle*, 1872). Heine n'est d'ailleurs pas le seul à avoir utilisé implicitement ce théorème, dans un cas particulier, longtemps avant qu'il ait été énoncé sous une forme générale. M. Lebesgue a utilisé fréquemment le théorème généralisé dans l'énoncé duquel on supprime le mot *dénombrable.* On a proposé de donner à ce théorème généralisé le nom de *théorème de Borel-Lebesgue. Voir* aux *C. R. Acad. Sc.* (t. CXLIV, 1907), les Notes de MM. Lebesgue et Schœnflies; l'article de M. Zoretti dans l'*Encyclopédie française* et le Livre de M. Paul Montel, *Sur les séries de polynomes.*

$a_n b_n$ et, à partir de ce moment, les n_k devraient nécessairement être pris inférieurs à n, c'est-à-dire seraient en nombre limité.

Il résulte évidemment du théorème fondamental que *s'il est possible d'enfermer tous les points d'un intervalle ab à l'intérieur d'intervalles $a_n b_n$, la longueur totale de ces intervalles est supérieure à la longueur de ab*. C'est cette conséquence du théorème fondamental qui est essentielle dans la théorie de la mesure des ensembles [1]. Si l'on observe qu'il suffit d'agrandir un intervalle d'une fraction aussi petite qu'on veut de sa longueur, pour que les extrémités de l'intervalle primitif soient intérieures (au sens étroit) à l'intervalle agrandi, on peut donner aussi l'énoncé suivant, parfois plus commode :

Si tout point d'un intervalle ab appartient à l'un au moins des intervalles $a_n b_n$, la somme des longueurs des $a_n b_n$ n'est pas inférieure à la longueur de ab.

Dans ce dernier énoncé, les extrémités d'un intervalle sont considérées comme appartenant à cet intervalle.

II. — *La mesure des ensembles bien définis.*

Lorsqu'un ensemble est formé d'un seul intervalle, sa mesure n'est pas autre chose que le rapport de sa longueur à la longueur choisie comme unité. Ce rapport est un nombre calculable, si les abscisses des extrémités de l'intervalle, évaluées avec l'unité choisie, sont elles-mêmes des nombres calculables [2].

Par définition, *la mesure d'un ensemble formé d'un nombre fini ou d'une infinité dénombrable d'intervalles sans points communs est égale à la somme de leurs mesures*. Il résulte du théorème fondamental que, si tous les points de l'ensemble sont intérieurs à un intervalle *ab*, la série qui définit la mesure est convergente et a une somme au plus égale à la mesure de *ab*. Nous choisirons, pour *ab* (*intervalle fondamental*, renfermant tous les ensembles que nous considérons), l'intervalle 0 — 1, de mesure égale à l'unité.

Il résulte de la définition, et des propriétés élémentaires des séries doubles

---

(1) Si l'on admettait les définitions idéalistes, c'est-à-dire si l'on considérait comme donné un ensemble *non dénombrable* d'intervalles, on pourrait observer que la somme des longueurs de ces intervalles dépasse forcément *ab*, et que par suite, en ce cas, la conséquence est vraie sans que le théorème fondamental soit nécessaire. C'est ainsi que j'avais été conduit, ayant en vue cette conséquence, à introduire l'hypothèse de l'énumérabilité des intervalles dans l'énoncé du théorème fondamental (*voir* ma Thèse : *Sur quelques points de la théorie des fonctions*, Paris, 1894).

(2) On pourrait concevoir un intervalle défini d'une manière telle que sa mesure soit calculable, sans que les abscisses de ses extrémités le soient; il suffit de supposer que ces abscisses sont x_1 et $x_1 + x_2$, x_1 n'étant pas calculable et x_2 étant calculable. Nous exclurons ce cas.

à termes positifs, que *la mesure d'un nombre fini ou d'une infinité dénombrable d'ensembles sans points communs est égale à la somme de leurs mesures.*

Les deux points précédents subsistent, lorsque les points communs sont en infinité dénombrable.

La mesure n'a été définie jusqu'ici que pour les ensembles formés d'intervalles (en nombre fini ou infini); si un tel ensemble E a pour mesure α, l'ensemble complémentaire par rapport à $0-1$ (l'ensemble des points de $0-1$, qui n'appartiennent pas à E) aura, par définition, pour mesure $1-\alpha$. Plus généralement, si deux ensembles, dont la mesure a été définie, sont tels que tous les points du second appartiennent au premier, *leur différence*, c'est-à-dire l'ensemble des points du premier qui n'appartiennent pas au second, *a pour mesure la différence des mesures.* La définition de la mesure d'une somme d'ensembles sans parties communes s'étend aux nouvelles mesures ainsi définies, et la définition de la mesure de la différence s'étend aussi à ces nouvelles mesures; on arrive ainsi à obtenir la mesure de tout ensemble *bien défini* par les procédés mêmes au moyen desquels l'ensemble a pu être bien défini.

J'ai indiqué cette marche dans mes *Leçons sur la théorie des fonctions;* comme j'avais surtout en vue les applications, je me suis contenté d'affirmer que le théorème fondamental, ou des procédés tout à fait analogues à ceux qu'on a employés pour démontrer ce théorème, permettent de justifier ces définitions en prouvant qu'elles ne sont jamais contradictoires entre elles. Mais j'ai omis toute démonstration, car la rédaction détaillée me paraissait devoir être longue et fastidieuse. Cette justification résulte indirectement des travaux de M. Lebesgue, publiés depuis, auxquels je pourrais renvoyer; mais il me semble préférable de développer complètement la théorie en restant au point de vue des définitions constructives; car c'est la forme sous laquelle j'ai été naturellement conduit à ces considérations, à propos de questions qui se sont posées dans mes recherches de théorie des fonctions (notamment dans ma *Thèse*) et c'est aussi la forme sous laquelle les questions se posent dans les applications. J'exposerai cette méthode *constructive* sous deux formes différentes, la première plus analytique et la seconde plus synthétique. La marche analytique est plus longue, mais me paraît être plus instructive; je vais tâcher de la résumer en conservant seulement les points essentiels.

Les deux opérations fondamentales, par lesquelles on construit des ensembles bien définis au moyen d'intervalles, consistent à faire la somme d'une infinité énumérable d'ensembles déjà définis (sans point commun) et à prendre la différence de deux ensembles (dont l'un contient l'autre). Ces opérations peuvent évidemment être répétées une infinité énumérable de fois; on peut, d'ailleurs, concevoir des opérations qui supposent qu'on ait effectué préalablement une infinité énumérable d'opérations préliminaires, et ainsi de suite. On sait que tous les processus de ce genre sont susceptibles d'une étude générale, qui a été faite pour la première fois par M. Georges Cantor, à l'occasion des dérivés successifs d'un ensemble donné,

et qui conduit à la notion des nombres transfinis de la seconde classe. Un nombre transfini de la seconde classe n'est pas autre chose qu'une notation abrégée, pour indiquer l'ordre dans lequel doivent être effectuées une infinité énumérable d'opérations, comportant une infinité énumérable de passages à la limite successifs ou superposés. Par exemple, la notation ω^2 désigne une infinité d'opérations $U_{n,p}$ correspondant aux couples d'entiers positifs n et p, et effectuées dans un ordre tel que l'opération $u_{n',p'}$ précède l'opération $u_{n,p}$ dans le cas où $n' < n$ et aussi dans le cas où $n' = n$, $p' < p$; l'opération $u_{1,o}$, par exemple, suppose donc effectuées préalablement les opérations $u_{o,p}$ quel que soit p, et fait intervenir les résultats de cette infinité d'opérations.

On doit considérer un nombre transfini comme bien défini, lorsque l'on connaît d'une manière précise l'ordre des opérations correspondantes; il est alors aisé d'indiquer une notation désignant ce nombre. Il est évident que l'ensemble des nombres transfinis qui peuvent être bien définis est dénombrable; mais il n'est pas effectivement énumérable; c'est une des formes du paradoxe du transfini, sur lequel je me suis expliqué ailleurs (1). On démontre habituellement que tout ensemble dénombrable bien ordonné définit un nombre transfini; mais il est aisé de voir que la démonstration suppose l'ensemble effectivement énumérable (2); en d'autres termes, il faut connaître une correspondance effective entre les éléments de l'ensemble et les entiers positifs; si l'on remplace chaque élément de l'ensemble par l'entier correspondant, ceci revient à ranger les entiers positifs sous la forme d'un ensemble bien ordonné, semblable à l'ensemble considéré, c'est-à-dire qu'on suppose la connaissance effective de la description des opérations de passage à la limite successives qui correspondent au nombre transfini, c'est-à-dire précisément la connaissance effective de ce nombre transfini.

Cette digression sur les nombres transfinis n'avait point d'autre but que de délimiter exactement la portée de la méthode employée, au point de vue adopté dans ce Mémoire; cette méthode s'étend évidemment à *tous* les nombres transfinis de la seconde classe, pour ceux qui attachent un sens à ces mots; d'autre part, dans l'enseignement élémentaire, l'exposition peut en être simplifiée en se bornant aux opérations d'ordre inférieur à ω^ω (ce qui correspond, dans la théorie des fonctions, aux fonctions de classe finie de M. Baire); cette limitation suffit dans la plupart des applications.

Le problème consiste maintenant, un ensemble bien défini étant construit au moyen d'intervalles par une série d'opérations correspondant à un nombre transfini déterminé, à évaluer la mesure de cet ensemble bien défin à partir de sa construction, et à montrer qu'on ne peut être conduit ainsi à aucune contradiction.

On est ainsi amené à faire la somme d'un certain nombre de séries ayant pour termes les longueurs des intervalles, puis de nouvelles séries ayant

(1) *Annales de l'École Normale*, 1908. (Ci-dessus, Note IV, § VI.)
(2) *Voir*, par exemple, Baire, *Leçons sur les fonctions discontinues*.

pour termes les sommes précédentes ou certaines de leurs différences deux à deux, et ainsi de suite (la notation du nombre transfini précisant ce que veut dire *ainsi de suite*). Bien entendu, les extrémités des intervalles sont définies par des nombres calculables; pour que les opérations qui donnent la mesure puissent être effectuées, il faut que toutes les séries qui interviennent soint convergentes [1]. De plus, pour que la définition de la mesure ne soit pas contradictoire, il est évidemment nécessaire qu'on soit assuré de trouver deux nombres égaux, si l'on définit la mesure d'un même ensemble E par deux procédés différents. Or, il résulte de la définition adoptée que, si un ensemble E a pour mesure α, l'ensemble complémentaire (par rapport à $0-1$) a pour mesure $1-\alpha$; si donc on était amené à attribuer à un ensemble E, par deux procédés différents, deux mesures différentes α et β, on serait amené à attribuer à l'ensemble $0-1$ les mesures $1-\alpha+\beta$ et $1-\beta+\alpha$, nombres dont l'un est supérieur à l'unité [2]. Il suffit donc de faire voir que la définition de la mesure ne peut pas conduire à attribuer à l'ensemble des points compris entre 0 et 1 une mesure supérieure à *un;* la démonstration de ce fait comprendra celle de la convergence des séries à termes positifs qui interviennent dans la définition.

Considérons donc un processus arbitraire, mais bien déterminé, conduisant à la mesure de l'ensemble $(0-1)$ par une série d'opérations correspondant à un nombre transfini donné, c'est-à-dire comportant une infinité dénombrable de passages à la limite effectués dans un ordre connu: ces passages à la limite sont donc effectivement énumérables; les opérations comportent, en outre, des soustractions qui ne sont pas des passages à la limite, mais qui sont en quelque sorte enchevêtrées d'une manière arbitraire (mais connue aussi) au milieu de ces passages à la limite. Si nous voulons effectuer le calcul de la mesure avec une approximation donnée, c'est-à-dire en commettant une erreur inférieure à un nombre donné ε, nous sommes conduits à nous donner une série à termes positifs de somme ε

$$\varepsilon = \varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_n + \ldots$$

et à effectuer les opérations de passage à la limite (en infinité énumérable) avec des erreurs respectivement inférieures à $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n, \ldots$; c'est cette marche que nous allons suivre pas à pas.

Nous rencontrons d'abord des ensembles formés par la réunion d'une infinité énumérable d'intervalles sans partie commune; la somme des lon-

---

(1) Il serait même nécessaire que la *rapidité* de la convergence en fût connue d'une manière précise; nous avons déjà observé que cette exigence devrait être formulée à propos de tous les calculs par séries convergentes; dans la pratique, on se contente habituellement de constater la convergence empiriquement.

(2) Les deux procédés de construction qui conduisent aux mesures α et β correspondent, en général, à deux nombres transfinis différents, dont l'un est supérieur à l'autre. Le procédé de construction qui conduirait à attribuer à $0-1$ une mesure supérieure à l'unité correspond au plus grand de ces deux nombres.

gueurs d'un nombre fini d'entre eux est évidemment inférieure à l'unité; ils forment donc une série convergente, qu'on calculera à ε_j près, en ne conservant qu'un nombre fini de termes. Si l'on prend ensuite la différence des mesures de deux ensembles, dont les mesures sont connues respectivement à ε_j près et à ε_k près, l'erreur sur la différence sera au plus en valeur absolue $\varepsilon_j + \varepsilon_k$.

Avant d'effectuer des opérations sur les ensembles ainsi obtenus par une première application de la méthode de construction, il est utile d'examiner d'un peu plus près la structure de ces ensembles. Considérons d'abord ceux qui sont obtenus par des sommations (sans soustraction).

Chacun d'eux peut être considéré comme formé par la réunion de deux parties : des intervalles en nombre fini (*partie fondamentale*); des intervalles en nombre infini, dont la somme des longueurs est inférieure à ε_j (*partie complémentaire*). Pour faire la différence de deux tels ensembles dont l'un contient l'autre, on est conduit à faire d'abord la différence des deux parties fondamentales. Il pourra se faire que la partie à retrancher, dont tous les points appartiennent par hypothèse à l'ensemble dont on retranche, n'appartienne pas tout entière à la partie fondamentale de cet ensemble; mais les points qui n'appartiennent pas à cette partie fondamentale appartiennent à la partie complémentaire; ils constituent un nombre fini d'intervalles de longueur inférieure à ε_j.

On est ainsi conduit à considérer, comme *la partie fondamentale de l'ensemble-différence*, l'ensemble formé d'un nombre fini d'intervalles qu'on obtient en faisant la différence des deux parties fondamentales, sans tenir compte, au cas où elles existeraient, des portions de la partie soustraite qui seraient à l'extérieur de la partie dont on soustrait. L'ensemble-différence est donc formé d'une partie fondamentale comprenant un nombre fini d'intervalles et d'une partie complémentaire de caractère nouveau, car elle comprend des intervalles additifs et des intervalles soustractifs; j'insiste un peu sur cette notion nouvelle, car elle est essentielle. Je considère la *somme* des parties complémentaires des ensembles dont on fait la différence (en ne comptant qu'une fois les parties communes), c'est un ensemble dénombrable d'intervalles dont la somme des longueurs est inférieure à $\varepsilon_j + \varepsilon_k$. L'ensemble-différence peut alors être défini comme il suit : il se compose de la partie fondamentale à laquelle on ajoute ou de laquelle on retranche, d'une manière qu'il ne sera pas nécessaire de préciser davantage, des points appartenant tous à cette partie complémentaire inférieure à $\varepsilon_j + \varepsilon_k$.

Considérons maintenant une infinité d'ensembles ainsi composés chacun d'une partie fondamentale (comprenant un nombre fini d'intervalles) et d'une partie complémentaire $\Sigma\varepsilon_j$ dont les points sont additifs ou soustractifs; les ε_j figurant dans les diverses parties complémentaires sont, bien entendu, tous distincts [1]; je dis tout d'abord que, si l'on considère une

---

[1] Il n'est pas interdit de faire figurer une infinité de fois un même ensemble dans la définition; mais si un même passage à la limite est ainsi répété une

infinité de tels ensembles n'ayant deux à deux aucun point commun, la somme des longueurs de leurs parties fondamentales est une série convergente. Ces parties fondamentales peuvent avoir des parties communes; mais ces parties communes sont nécessairement intérieures [1] aux ε_j; la somme des longueurs d'un nombre fini des parties fondamentales est donc au plus égale à l'unité, augmentée de la somme des ε_j: la série est, par suite, convergente; nous avons, par hypothèse, fait correspondre à l'opération de passage à la limite qu'est la sommation de cette série convergente, un ε, soit ε_k; nous négligerons le reste de cette série convergente en prenant un nombre de termes tels que ce reste soit inférieur à ε_k, et la partie fondamentale de l'ensemble obtenu sera, par définition, la réunion des parties fondamentales, en nombre fini, qui subsistent lorsqu'on néglige ce reste; les autres parties fondamentales seront jointes à l'ensemble des parties complémentaires.

Nous obtenons donc, par un nouveau passage à la limite, un ensemble de même structure que ceux dont nous sommes partis; il est formé d'une partie fondamentale, comprenant un nombre fini d'intervalles, et d'une partie complémentaire consistant en intervalles complémentaires $\Sigma\varepsilon_j$, dont les points sont additifs ou soustractifs. Il n'est pas besoin d'insister : la répétition des mêmes opérations conduira toujours à des résultats analogues; on verra simplement apparaître, à chaque nouveau passage à la limite, un nouveau terme dans $\Sigma\varepsilon_j$, mais cette somme sera toujours inférieure à ε. En définitive, l'ensemble que l'on considère (et qui est actuellement l'ensemble $0-1$) est formé d'une partie fondamentale et d'une partie complémentaire inférieure à ε. Du moment que tous les points de l'ensemble $0-1$ sont intérieurs, soit aux intervalles de la partie fondamentale, soit aux intervalles de la partie complémentaire, la somme des longueurs de tous ces intervalles est, d'après le théorème fondamental. supérieure à 1; donc, la mesure de la partie fondamentale est comprise entre $1-\varepsilon$ et 1.

Mais cette mesure peut aussi être obtenue de proche en proche par les opérations au moyen desquelles on a construit l'ensemble; la seule difficulté provient de ce que, lorsqu'on réunit des ensembles sans points communs, il peut arriver que leurs parties fondamentales aient des parties communes, si ces parties communes appartiennent par ailleurs à des intervalles d'exclusion négatifs. Mais il suffit, pour écarter cette difficulté, d'observer que les portions parasites ainsi introduites étant certains intervalles complémentaires, leur somme algébrique (car elles peuvent devenir négatives dans les opérations de soustraction) est, au plus, égale à $\Sigma\varepsilon_j$, c'est-

---

infinité de fois, chacune de ces opérations occupe un rang distinct parmi l'infinité énumérable de toutes les opérations considérées, et il lui correspond par suite, chaque fois, un ε_j d'indice différent.

[1] Il peut même arriver que les parties fondamentales aient en commun plusieurs fois un même intervalle; il doit alors appartenir à plusieurs des ε_j distincts.

à-dire à ε. Le nombre auquel on est conduit, pour la mesure, diffère donc de moins de ε de la mesure de la partie fondamentale : nous avons vu que cette dernière mesure est comprise entre $1-\varepsilon$ et 1; la mesure de l'ensemble $0-1$, par le procédé considéré, est donc comprise entre $1-2\varepsilon$ et $1+\varepsilon$; mais cette mesure est un nombre fixe et ε est arbitraire; ce nombre fixe est donc rigoureusement égal à 1, ce que nous voulions démontrer.

J'ai tenu à exposer d'abord la marche analytique qui me paraît la plus naturelle, et qui a l'avantage de devenir extrêmement simple et élémentaire lorsqu'on se borne à considérer les nombres transfinis inférieurs à ω^ω. L'exposition se trouve en ce cas très simplifiée et met nettement en évidence la structure des ensembles formés.

Lorsqu'on ne limite pas les nombres transfinis qu'on introduit, l'exposition synthétique que je vais maintenant indiquer est plus brève; la méthode est la même et consiste toujours à déduire les ensembles nouveaux d'ensembles déjà définis; seulement, au lieu de partir des intervalles et de suivre la construction de proche en proche, on suppose que la construction a été faite jusqu'à un certain point et possède certaines propriétés, et l'on démontre que ces propriétés subsistent lorsqu'on avance d'un pas nouveau.

Nous donnerons le nom de *corps ouvert* d'ensembles à une infinité d'ensembles A ayant les propriétés suivantes :

1° Chaque ensemble A se déduit d'autres ensembles A au moyen des opérations fondamentales (addition d'une infinité énumérable d'ensembles A sans point commun; différence de deux ensembles A dont l'un contient l'autre); on peut ainsi, de proche en proche, ramener chaque ensemble A à être défini à partir des ensembles élémentaires (intervalles).

2° Quel que soit le nombre positif ε, chaque ensemble A peut être regardé comme formé d'une partie fondamentale (nombre fini d'intervalles), à des ensembles près positifs ou négatifs enfermés dans des intervalles d'étendue totale inférieure à ε.

3° On peut faire correspondre à chaque ensemble A un nombre, qu'on appelle *sa mesure*, et qui se déduit des mesures des intervalles par les mêmes opérations constructives au moyen desquelles l'ensemble A se déduit des intervalles (la mesure d'un intervalle est égale à sa longueur). La mesure de la partie fondamentale d'un ensemble A tend vers la mesure de A lorsque ε tend vers zéro. Cette propriété entraîne comme conséquence que, si l'on peut obtenir A de plusieurs manières différentes à partir des intervalles, la valeur de la mesure déduite de ces diverses constructions est la même.

Cela posé, on a le théorème suivant :

Si l'on adjoint à un corps ouvert la différence de deux ensembles A *de ce corps dont l'un contient l'autre, ou la somme d'une infinité dénombrable d'ensembles* A *du corps (sans parties communes), on*

obtient un autre corps ouvert (¹) *ayant les mêmes propriétés*, ce qu'on peut exprimer brièvement en disant que les nouveaux ensembles obtenus sont aussi des ensembles A.

Il est évident que la différence de deux ensembles A est un ensemble A; je dis qu'il en est de même de la somme d'une infinité dénombrable d'ensembles $A_1, A_2, \ldots, A_n, \ldots$, sans parties communes. Il suffira, en effet, de faire correspondre à chaque A_n un nombre ε_n tel que la série $\Sigma\varepsilon_n$ soit convergente. Soit E_n la partie fondamentale de A_n; la série des mesures des E_n est convergente (car les parties communes aux E_n sont des intervalles dont la somme est, au plus, $\Sigma\varepsilon_n$); on pourra donc choisir n tel que la somme des mesures des E_{n+k} $(k > 0)$ soit inférieure à ε. La réunion de $E_1, E_2, \ldots, E_n$ forme un ensemble analogue E (formé d'un nombre fini d'intervalles); l'ensemble A, réunion des $A_1, A_2, \ldots, A_n, \ldots$, ne diffère donc de E que par des ensembles complémentaires enfermés dans des intervalles au plus égaux à $\varepsilon + \Sigma\varepsilon_n$.

Je dis, de plus, que si l'on a démontré que la mesure de la partie fondamentale de A_n est aussi voisine qu'on veut de la mesure de A_n, la même propriété subsiste pour A. En effet, la mesure de A est, par définition, la somme des mesures des A_n sans parties communes; elle diffère donc aussi peu qu'on veut de la somme des mesures des E_n; mais cette dernière somme ne diffère de la mesure de E que de $\varepsilon + \Sigma\varepsilon_n$, au plus; la différence entre la mesure de E et la mesure de A peut donc être rendue aussi petite qu'on veut La mesure de A est, par suite, indépendante du procédé particulier par lequel A a été obtenu; elle serait la même, si A était construit au moyen d'autres ensembles du corps ouvert primitif.

III. — *Le second théorème fondamental.*

Les considérations précédentes s'étendent, sans difficulté, au cas de n dimensions (²). Je vais les énoncer en précisant les définitions.

(¹) Le *paradoxe du transfini* consiste précisément en ce que, par la répétition indéfinie de ce procédé, on n'obtiendra jamais un corps *fermé*, c'est-à-dire ne pouvant plus être étendu par la répétition des mêmes opérations. Pour les raisons que j'ai déjà exposées ailleurs (*Annales de l'École Normale*, 1908) et sur lesquelles je ne reviens pas, la conception d'un tel *corps fermé*, bien que n'étant pas contradictoire en soi, ne me paraît pas être une véritable conception mathématique, parce qu'il n'est pas possible de décrire exactement, par un nombre fini de mots, la construction d'un tel corps. D'autre part, on n'aura jamais besoin, dans aucune application, que d'un des corps ouverts que nous avons définis. Pour les mathématiciens qui admettent l'existence de *tous* les nombres transfinis de seconde classe, les considérations du texte permettent, par l'application transfinie du même procédé, d'obtenir *un corps fermé* ayant les mêmes propriétés que les *corps ouverts*.

(²) L'extension au cas d'un nombre infini de dimensions est aisée avec les hypothèses de convergence bien connues.

J'appelle *ensemble élémentaire* l'ensemble des points dont les n coordonnées $x_1, x_2, \ldots, x_n$ vérifient les inégalités

$$a_1 \leqq x_1 \leqq b_1; \quad \ldots, \quad a_n \leqq x_n \leqq b_n.$$

Sa mesure est le produit

$$(b_1 - a_1)(b_2 - a_2)\ldots(b_n - a_n),$$

la réunion d'un nombre fini d'*ensembles élémentaires* constitue un *domaine simple.*

Les ensembles *bien définis* sont les ensembles qui peuvent être obtenus, à partir des ensembles élémentaires, au moyen des deux opérations fondamentales suivantes indéfiniment répétées :

1° Faire la différence des deux ensembles A et B déjà définis et tels que tout point de B appartienne à A,

$$\text{(D)} \qquad A - B;$$

2° Faire la somme d'une infinité d'ensembles déjà définis : $A_1, A_2, \ldots, A_n, \ldots$, sans parties communes

$$\text{(S)} \qquad A_1 + A_2 + \ldots + A_n + \ldots.$$

On peut alors énoncer le théorème suivant, que nous appellerons *second théorème fondamental :*

Étant donnés, dans un domaine borné, un ensemble bien défini quelconque E *et un nombre arbitraire* ε, *on peut trouver un domaine simple* D *tel que l'ensemble des points de* E, *qui n'appartiennent pas à* D, *et l'ensemble des points de* D *qui n'appartiennent pas à* E *puissent être enfermés dans une infinité dénombrable d'ensembles élémentaires, dont la somme des mesures soit inférieure à* ε.

Plus brièvement, *tout ensemble* E *équivaut, à* ε *près, à un domaine simple* D.

Il est quelquefois commode d'adjoindre aux opérations (D) et (S) l'opération (P), qui consiste à prendre la partie commune aux ensembles A et B

$$\text{(P)} \qquad (A, B).$$

Si l'on part du *corps* constitué par l'ensemble des intervalles (qu'on peut supposer à extrémités calculables, pour préciser), il est clair que le résultat de l'opération (P) effectuée sur deux éléments A et B du corps, c'est-à-dire sur deux intervalles, peut être obtenu en appliquant les opérations (D) et (S) à d'autres éléments du corps. Je dis que si cette propriété est vraie pour un corps, elle subsiste lorsqu'on étend ce corps en lui adjoignant les résultats des opérations (D) ou (S) effectuées sur ces

éléments. Cela est à peu près évident en ce qui concerne (D); détaillons le calcul pour (S); soit

$$A = A_1 + A_2 + \ldots + A_n + \ldots,$$
$$B = B_1 + B_2 + \ldots + B_p + \ldots.$$

On a évidemment

$$(A, B) = \Sigma\Sigma(A_n, B_p).$$

On peut donc adjoindre l'opération (P) aux opérations (D) et (S); on obtiendra toujours des ensembles *bien définis*, mais la mesure de ces ensembles résultera moins directement de leur définition.

IV. — *Les ensembles à points multiples.*

On sait que si l'on étend à un contour plan fermé C l'intégrale curviligne

$$J = \frac{1}{2}\int_C x\,dy - y\,dx,$$

on obtient, si C est un contour simple, la valeur, au signe près, de l'aire intérieure à C. Si le contour C a des boucles, la valeur de l'intégrale J est, en général,

$$J = nS_n + (n-1)S_{n-1} + \ldots + S_1 - S'_1 - 2S'_2 - \ldots - pS'_p,$$

$S_n, S_{n-1}, \ldots, S_1, S'_1, \ldots, S'_p$ étant les sommes des aires de certains des domaines déterminés dans le plan par la courbe C. Chacun des points intérieurs à S_n est ainsi compté n fois.

On est ainsi naturellement conduit à envisager des ensembles, dont chaque point est affecté d'un nombre entier, représentant le nombre de fois que ce point figure dans l'ensemble. Nous nous bornerons au cas où ces nombres entiers sont tous positifs et nous les supposerons de plus, tout d'abord, inférieurs à un nombre fixe N.

Je signale en passant qu'on pourrait, en passant du cas des entiers au cas des nombres commensurables, puis incommensurables, déduire de ceci une théorie élémentaire de l'intégrale définie, au sens de M. Lebesgue; mais je ne m'y attarderai pas. Je vais d'ailleurs me borner, pour simplifier le langage, aux ensembles linéaires intérieurs à $0-1$.

Un tel ensemble sera obtenu, comme précédemment, par la réunion d'intervalles; mais les parties communes à deux ou plusieurs intervalles seront considérées comme comptant deux ou plusieurs fois. Nous continuerons à utiliser les opérations (D) et (S); pour ne pas introduire de nombres négatifs, nous ne considérerons que la différence A — B d'ensembles A et B tels que tout point de B figure dans A à un degré de multiplicité au moins égal à celui qu'il a dans B; pour la somme, nous supposerons qu'il peut y avoir des parties communes, le degré de multiplicité de chaque point restant cependant inférieur à N.

Il est manifeste que l'ensemble des points de degré de multiplicité k est *bien défini;* si l'on désigne par σ_k sa mesure, la mesure de l'ensemble considéré est, par définition,

$$\sigma = \sum_1^N k\,\sigma_k.$$

Or, on a évidemment

$$\Sigma\sigma_k \leqq 1.$$

Il en résulte donc

$$\sigma \leqq N.$$

Si l'on avait $\Sigma\sigma_k < s$, il en résulterait $\sigma \leqq Ns$.

Cette inégalité conduit très aisément à un théorème que j'ai énoncé, il y a plusieurs années (¹), et dont je n'avais pas encore publié la démonstration.

Si l'on a dans l'intervalle $0 - 1$ *une infinité dénombrable d'ensembles* (bien définis) E_n *tels que la mesure de chacun d'eux soit supérieure ou égale à* σ, *l'ensemble* E *des points communs à une infinité d'entre eux a une mesure qui n'est pas inférieure à* σ.

L'ensemble E est évidemment bien défini; nous allons montrer que l'hypothèse où sa mesure σ' serait inférieure à σ conduit à une contradiction. Soient, en effet, e_0 l'ensemble des points qui n'appartiennent à aucun des E_i, e_k l'ensemble des points qui appartiennent à k d'entre eux (et n'appartenant pas à $k+1$); il est clair que tout point de $0 - 1$ appartient à l'un et à un seul des ensembles E, e_0, e_1, ..., e_k, ...; si donc on désigne par σ_k la mesure de e_k, on a

$$\sigma' + \sigma_0 + \sigma_1 + \ldots + \sigma_k + \ldots = 1,$$

la série étant forcément convergente.

Le nombre σ' étant, par hypothèse, inférieur à σ, on peut choisir n assez grand pour qu'on ait

$$\sigma_{n+1} + \sigma_{n+2} + \ldots < \frac{\sigma - \sigma'}{2},$$

et il en résulte

$$\sigma' + \sigma_{n+1} + \sigma_{n+2} + \ldots < \sigma' + \frac{\sigma - \sigma'}{2}.$$

Ceci posé, considérons l'ensemble à points multiples $\mathcal{E}$ formé par la réunion de E_1, E_2, ..., E_N; sa mesure est supérieure ou égale à $N\sigma$, puisque la mesure de chacun des E_i est supérieure ou égale à σ. Nous allons évaluer cette mesure par une autre voie. Il est clair que l'ordre de multiplicité d'un point de $\mathcal{E}$ ne peut dépasser N et que cet ordre est, d'autre part, au plus égal à l'ordre de multiplicité du même point dans l'ensemble de *tous* les ensembles E_i (ce dernier ordre pouvant être infini). On aura donc une

(¹) *C. R. Acad. Sc.*, 7 décembre 1903.

limite supérieure de la mesure de $\mathcal{E}$, en la calculant comme si tous les points de e_k avaient dans $\mathcal{E}$ un ordre de multiplicité égal à k lorsque k est inférieur à N, et égal à N lorsque k est égal ou supérieur à N, les points de E (d'ordre infini) étant aussi comptés comme d'ordre N dans $\mathcal{E}$. On trouve ainsi que la mesure de $\mathcal{E}$ est au plus égale à

$$\sigma_1 + 2\sigma_2 + \ldots + (N-1)\sigma_{N-1} + N(\sigma_N + \sigma_{N+1} + \ldots + \sigma');$$

mais, si N est supérieur à n, ceci est inférieur à

$$n(\sigma_1 + \sigma_2 + \ldots + \sigma_n) + N(\sigma_{n+1} + \sigma_{n+2} + \ldots + \sigma'),$$

c'est-à-dire à

$$n + N\left(\sigma' + \frac{\sigma - \sigma'}{2}\right).$$

Or, n et σ étant fixes, on peut toujours prendre N assez grand pour que cette dernière expression soit inférieure à $N\sigma$; l'hypothèse $\sigma' < \sigma$ conduit donc à une contradiction.

Ce théorème conduit fort simplement à une propriété des séries, qui a dû être souvent rencontrée par ceux qui se sont occupés de ces questions, et dont j'ai indiqué une application à la théorie des fonctions, dans la Note que je viens de citer.

Soient une série convergente pour toutes les valeurs de x comprises entre 0 et 1, et ε_n un nombre positif quelconque. Étant donné un nombre positif $\sigma < 1$, on peut évidemment déterminer un ensemble E_n, de mesure supérieure à σ et tel que, pour tous les points de E_n, il existe un nombre N, tel que les restes de la série de rang égal ou supérieur à N soient inférieurs à ε_n pour tous les points de E_n. En effet, à tout point x correspond un nombre ν, tel que les restes de rang égal ou supérieur à ν soient inférieurs à ε_n; l'ensemble des valeurs de x, correspondant à un nombre donné ν, forment un ensemble e_ν; la somme des mesures de ces ensembles est convergente et égale à 1; on peut donc prendre N assez grand pour que le reste de cette série convergente soit inférieur à $1 - \sigma$.

Ceci étant, donnons-nous des nombres ε_n tendant vers 0; à chacun d'eux correspond un ensemble E_n de mesure supérieure à σ; l'ensemble E des points communs à une infinité des E_n a une mesure supérieure à σ; sur l'ensemble E, la série est évidemment uniformément convergente; c'est la proposition dont je voulais parler. Nous en rencontrerons bientôt, par une autre voie encore plus simple, un cas particulier, dont l'importance me paraît surpasser celle de la proposition générale.

V. — *Les ensembles de mesure nulle.*

D'après ce qui précède, un ensemble bien défini est de mesure nulle, lorsqu'il peut être enfermé, quelle que soit ε, à l'intérieur d'ensembles élémentaires de mesure totale inférieure à ε. Inversement, tout ensemble

qui a cette propriété fait partie d'un ensemble bien défini de mesure nulle; on ne peut donc lui attribuer une mesure autre que zéro. Cette remarque montre que, si l'on admet qu'il soit possible de considérer des ensembles autres que les ensembles bien définis, il peut exister des ensembles qui soient mesurables, en vertu de mes définitions, sans être bien définis; ce sont les ensembles qu'on obtient en ajoutant à un ensemble bien défini (de mesure non nulle) une portion *arbitraire* d'un ensemble bien défini de mesure nulle. La classe d'ensembles mesurables ainsi définie est équivalente à la classe d'ensembles que M. Lebesgue appelle *mesurables;* cette remarque a déjà été faite par M. Lebesgue; si je la rappelle, c'est pour éviter une confusion dont je suis responsable, car elle est la conséquence d'un langage mal choisi; j'avais appelé *mesurables*, dans mes *Leçons sur la théorie des fonctions*, les ensembles que j'appelle maintenant *bien définis;* M. Lebesgue a donné à ces ensembles le nom de *mesurables* B et l'on a cru parfois que je n'avais pas défini la mesure pour les ensembles que je n'appelais pas mesurables, bien que j'indique dans ce même Ouvrage que, si un ensemble est complètement intérieur à un ensemble mesurable, on devra regarder sa mesure comme inférieure ou égale à celle de l'ensemble mesurable, sans s'inquiéter s'il est mesurable ou non. La mesure ainsi définie sera connue avec précision dans les cas où l'on pourra démontrer à la fois qu'elle est supérieure ou égale et qu'elle est inférieure ou égale à un même nombre; c'est, en particulier, le cas pour les ensembles faisant partie d'un ensemble bien défini de mesure nulle; c'est aussi le cas pour tous les ensembles mesurables de M. Lebesgue. Mais je n'insiste pas sur ce point, tenant à me borner aux ensembles bien définis.

Tout ensemble linéaire de mesure nulle est intérieur à l'un des ensembles de mesure nulle, que j'ai nommés *réguliers* et qui sont définis au moyen d'une infinité dénombrable de points fondamentaux $A_1, A_2, \ldots, A_n, \ldots$. On définit d'abord un ensemble E_k au moyen d'intervalles entourant chacun des points fondamentaux et dont la somme σ_k est convergente; lorsque k augmente, l'intervalle attaché à un point fondamental quelconque A_n tend régulièrement vers zéro, sans jamais croître; lorsque σ_k tend vers zéro, l'ensemble E_k a pour limite, par définition, un ensemble régulier de mesure nulle. Je n'insisterai pas sur cette théorie, sur laquelle j'ai donné des indications suffisantes dans une Note des *Comptes rendus de l'Académie des Sciences* (¹); l'extension de la théorie de la similitude aux ensembles à plusieurs dimensions appellerait d'ailleurs de nouvelles recherches. Il serait aussi du plus haut intérêt d'approfondir, sans complications inutiles, la théorie des ensembles à plusieurs dimensions, dont les projections sont de mesure nulle, par exemple des ensembles situés dans l'espace, dont les projections sur un plan quelconque sont de mesure nulle. Enfin, je puis signaler comme se rattachant au même ordre d'idées, la

(¹) Séance du 6 mars 1911.

classification des ensembles de mesure nulle d'après la loi de décroissance asymptotique des intervalles au moyen desquels on peut les définir ([1]). Cette classification me paraît devoir être importante dans beaucoup de questions; mais je ne l'utiliserai pas ici.

CHAPITRE II.

La définition de l'intégrale.

Je me bornerai, bien entendu, à la considération des fonctions calculables, avec une restriction cependant, qui est capitale : si une fonction coïncide avec une fonction calculable presque partout, c'est-à-dire en tous les points qui n'appartiennent pas à un certain ensemble de mesure nulle, tout se passera, au point de vue de ses propriétés moyennes, comme si elle était calculable partout. En fait, nous constaterons qu'on peut ramener l'intégration des fonctions calculables à celle des polynomes, qui sont évidemment les plus simples des fonctions calculables.

I. — *La définition des fonctions bornées et le troisième théorème fondamental.*

Pour définir une fonction, on la considère généralement comme la limite d'une suite de fonctions connues; si l'on part des polynomes, considérés comme connus, on définira ainsi de proche en proche des fonctions de plus en plus compliquées. Un théorème de Weierstrass, dont nous n'aurons pas besoin, apprend que toute fonction continue peut être regardée comme limite vers laquelle tend *uniformément* une suite convergente de polynomes ([2]); ceci prouverait, s'il en était besoin, que les fonctions continues et les limites de fonctions continues et les limites de ces limites, etc. rentrent comme cas particuliers dans l'ensemble des fonctions que nous considérons. Mais il est évident, sans qu'il soit besoin du théorème de Weierstrass, que cet ensemble comprend toutes les fonctions définissables analytiquement, puisque tout procédé de définition analytique se ramène à un processus limite portant sur des fonctions antérieurement définies; et l'on doit remonter ainsi de proche en proche jusqu'à ce qu'on arrive à des fonctions connues; or, les fonctions élémentaires, définies directement (comme les fonctions circulaires par exemple), sont développables en séries entières et sont, par suite, des limites de polynomes.

Considérons donc une fonction $f(x, y, z)$, que nous supposons bornée dans un domaine D; nous dirons que *cette fonction est asymptotique-*

([1]) *C. R. Acad. Sc.*, 26 février 1912 : *La classification des ensembles de mesure nulle et la théorie des fonctions monogènes uniformes.* Voir mon Livre *Sur les fonctions monogènes uniformes d'une variable complexe.*

([2]) Voir mes *Leçons sur les fonctions des variables réelles*, Chap. IV.

ment équivalente à des polynomes, si à tout groupe de nombres positifs ε *et* α *on peut faire correspondre un polynome*

$$P(x, y, z; \varepsilon, \alpha),$$

tel que l'ensemble des points où la valeur absolue de la différence $f - P$ *est supérieure à* ε, *soit de mesure inférieure à* α, *c'est-à-dire soit intérieur à une infinité dénombrable d'ensembles élémentaires d'étendue totale inférieure à* α. Si de plus la fonction est bornée, nous supposerons les polynomes bornés dans leur ensemble.

Le théorème fondamental de la théorie des fonctions bornées est alors le suivant :

TROISIÈME THÉORÈME FONDAMENTAL. — *La limite supposée bornée d'une suite de fonctions bornées asymptotiquement équivalentes à des polynomes est elle-même asymptotiquement équivalente à des polynomes.* En d'autres termes, *toute fonction* [(*bornée*) (¹)] *définissable analytiquement est asymptotiquement équivalente à des polynomes.* Soient, en effet, $f_1, f_2, \ldots, f_n, \ldots$ des fonctions données; $\alpha_1, \alpha_2, \ldots, \alpha_n, \ldots$ des nombres positifs, dont la somme $\alpha = \Sigma\alpha_n$ est suffisamment petite; $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n, \ldots$ des nombres positifs décroissants et tendant vers zéro. Nous supposons que les fonctions f_n tendent vers une limite f en tout point de l'intervalle dans lequel elles sont définies. La fonction f étant bornée, on peut supposer que les fonctions f_n sont bornées dans leur ensemble. Sinon, on remplacerait par o les valeurs que prend chacune d'elles, lorsqu'elles sont comprises en dehors de l'intervalle $-M-\varepsilon, M+\varepsilon$, si f est compris entre $-M, +M$; cette opération ne peut, en effet, modifier leur limite; nous raisonnerons sur les nouvelles fonctions f_n qui sont aussi asymptotiquement équivalentes à des polynomes.

A la fonction f_n et aux nombres ε_n et α_n correspond un polynome P_n, tel que la valeur absolue de $f_n - P_n$ soit inférieure à ε_n, sauf peut-être dans un ensemble de mesure α_n. Si donc, on exclut du domaine donné D l'ensemble de mesure au plus égale à α formé par la réunion de ces ensembles de mesure au plus égale à α_n, dans le domaine restant D', la limite des P_n existera et sera égale à la limite des f_n.

Les polynomes P_n tendant vers une limite, à tout nombre η et à tout point $M(x, y, z)$ de D' correspond un nombre $N(\eta, M)$, tel que les relations

$$n \geqq N, \qquad n' \geqq N$$

entraînent

$$|P_n(x, y, z) - P_{n'}(x, y, z)| < \eta. \tag{1}$$

Lorsque η est donné, il correspondra en général des nombres N diffé-

---

(¹) *Voir*, plus bas (§ IV), comment les fonctions finies non bornées peuvent être définies comme limites de fonctions bornées.

rents aux différents points M de D, mais à tout point M correspond une valeur finie de N.

Je désigne par $\Delta(N')$ l'ensemble des points M pour lesquels N peut être pris inférieur à N', cet ensemble diffère aussi peu qu'on veut de D', lorsque N' augmente indéfiniment; s'il n'en était pas ainsi, c'est-à-dire si $\Delta(N')$ n'avait pas pour limite D', lorsque N' augmente indéfiniment, il y aurait dans D' des points n'appartenant à aucun des $\Delta(N')$, ce qui est absurde. On peut donc, étant donné un nombre arbitrairement petit β, faire correspondre à η un nombre N', tel que $\Delta(N')$ diffère de D' de moins de β.

Si nous nous donnons les nombres arbitrairement petits $\alpha+\beta$ et $\eta+\varepsilon_1$, ce qui précède nous montre que le polynome $P_N(x, y, z)$ diffère de moins de $\eta+\varepsilon_1$ de la fonction f, limite de f_n, sauf peut-être dans un ensemble de mesure au plus égale à $\alpha+\beta$. La fonction limite des f_n est donc asymptotiquement équivalente à des polynomes [1].

Remarque. — Chaque inégalité (1) définit un domaine algébrique (un nombre limité d'intervalles dans le cas d'une variable), qui peut être remplacé, si l'on veut, par un domaine simple avec une erreur aussi petite qu'on veut. Les ensembles, dont la mesure intervient dans la démonstration, s'obtiennent au moyen de ces domaines simples; le calcul de leurs mesures ne présente aucune difficulté.

II. — *La définition de l'intégrale des fonctions bornées.*

Le troisième théorème fondamental permet de définir d'une manière très simple l'intégrale des fonctions bornées, cette définition fournissant en même temps un procédé de calcul effectif. La marche suivie est exactement la même que pour la mesure des ensembles (Chap. I); les détails dans lesquels nous sommes entrés à cette occasion nous permettront d'être ici plus brefs.

Je suppose connue la définition élémentaire de l'intégration des polynomes et ses propriétés principales, notamment le premier théorème de la moyenne. J'appelle *intégrale élémentaire d'un polynome* l'expression

$$\int_a^x \int_b^y \int_c^z P(x, y, z)\, dx\, dy\, dz.$$

L'intégrale définie étendue à un domaine simple s'exprime au moyen d'un nombre fini d'intégrales élémentaires. Occupons-nous d'abord des intégrales élémentaires des fonctions bornées.

---

(1) On peut remarquer que, si l'on n'imposait pas aux polynomes la condition d'être bornés dans leur ensemble, la démonstration exigerait seulement que la limite existe, c'est-à-dire que f soit finie, mais non nécessairement que f soit bornée; elle subsisterait pourvu que l'ensemble des points singuliers, où les f_n seraient infinis et où la limite n'existerait pas, soit de mesure nulle.

Considérons une suite de polynomes $P_n(x, y, z)$ asymptotiquement convergente dans un domaine D se réduisant à un *ensemble élémentaire* (parallélépipède rectangle), et supposons, en outre, que l'ensemble des polynomes P_n soit borné dans D, c'est-à-dire qu'il existe un nombre M, tel qu'on ait, quel que soit n et quels que soient x, y, z dans D,

$$|P_n(x, y, z)| < M.$$

Je dis que la suite des intégrales des P_n est uniformément convergente dans D; nous poserons, a, b, c étant le point de D dont les coordonnées sont les plus petites,

$$Q_n(x, y, z) = \int_a^x \int_b^y \int_c^z P_n(x, y, z)\, dx\, dy\, dz.$$

En effet, étant donnés les nombres ε et α, on peut déterminer N, tel que les inégalités

$$n \geqq N, \qquad n' \geqq N$$

entraînent

$$|P_n - P_{n'}| < \varepsilon,$$

sauf au plus dans un ensemble de mesure α.

On a, pour ces mêmes valeurs de n et n',

$$|Q_n - Q_{n'}| < \int_a^x \int_b^y \int_c^z |P_n - P_{n'}|\, dx\, dy\, dz < \varepsilon V + 2\alpha M,$$

V étant le volume de D. Les nombres V et M étant fixes, ε et α arbitrairement petits, la proposition est démontrée. Cette proposition légitime la *définition* suivante : *L'intégrale d'une fonction bornée dans un domaine élémentaire est égale à la limite des intégrales des termes d'une suite bornée de polynomes asymptotiquement convergente vers cette fonction.* Il est clair en effet que, si deux suites bornées de polynomes convergent asymptotiquement vers une même fonction, l'ensemble des deux suites (suite obtenue en intercalant les termes successifs de l'une entre les termes de l'autre) est aussi une suite asymptotiquement convergente (car la réunion de deux ensembles de mesure inférieure à ε forme un ensemble de mesure inférieure à 2ε).

Il résulte du théorème de la moyenne et de notre second théorème fondamental que l'intégrale d'une fonction bornée, étendue à un ensemble bien défini quelconque, peut être définie comme égale à la limite de l'intégrale correspondant à un domaine simple, tendant vers cet ensemble bien défini. Car la valeur absolue de l'intégrale étendue à un ensemble élémentaire est au plus égale au produit de la mesure de cet ensemble par une limite supérieure de la valeur absolue de la fonction dans le domaine total considéré. Cette définition ne peut donc conduire à aucune contradiction.

On peut arriver par une autre voie à la définition de l'intégrale d'une fonction bornée dans un domaine D non élémentaire; il suffit de rem-

placer les polynomes par d'autres polynomes asymptotiquement équivalents aux premiers dans D et asymptotiquement équivalents à zéro dans le complémentaire de D, par rapport à un domaine parallélépipédique Δ contenant D.

Étant donnés des polynomes P_n et un domaine D, on détermine aisément des polynomes Π_n différant aussi peu qu'on veut des P_n dans D et différant aussi peu qu'on veut de zéro dans la portion de Δ qui n'appartient pas à D (les régions dans lesquelles ces inégalités ne seraient pas vérifiées ayant des mesures inférieures à des nombres ε_n tendant vers zéro avec n).

En résumé, dans le cas d'une variable, toute fonction bornée définissable $f(x)$ est asymptotiquement équivalente à une suite de polynomes $P_n(x)$, suite qui est évidemment asymptotiquement convergente; par définition, l'intégrale indéfinie de $f(x)$ est la fonction continue définie par la suite uniformément convergente des $Q_n(x)$, intégrales des $P_n(x)$. On passe immédiatement de l'intégrale indéfinie à l'intégrale définie dans un domaine simple, puis dans un domaine quelconque au moyen du second théorème fondamental.

La méthode des approximations successives permet de ramener au calcul d'intégrales définies le calcul des solutions des équations différentielles, des équations intégrales et intégro-différentielles; si ces opérations renferment des fonctions bornées bien définies quelconques, on pourra les remplacer par une suite asymptotiquement équivalente de polynomes, c'est-à-dire, lorsqu'on aura fixé une limite de l'approximation qu'on désire obtenir, effectuer seulement les calculs sur des polynomes (1).

III. — *Les propriétés de l'intégrale des fonctions bornées.*

Il est manifeste que l'intégrale des fonctions bornées, telle que nous l'avons définie, possède les propriétés les plus importantes de l'intégrale

---

(1) Dans l'intervalle qui s'est écoulé entre la rédaction de ce Mémoire et son impression (dans le *Journal* de M. Jordan, 1912) ont paru, dans les *Comptes rendus de l'Académie des Sciences*, deux Notes sur l'intégration : *Sur quelques points de la théorie des fonctions sommables* par M. Frédéric Riesz (4 mars 1912) et *Sur les propriétés des fonctions mesurables* par M. N. Lusin (17 juin 1912). On trouvera dans ces Notes des renseignements bibliographiques intéressants sur des travaux antérieurs de MM. Weyl, F. Riesz, Egoroff et Lusin. Ce qui me paraît distinguer le point de vue de ces auteurs de celui que j'ai adopté, c'est que la définition de ce que j'appelle la *convergence asymptotique des suites de polynomes* ne repose que sur la mesure des domaines *algébriques* (dans le cas d'une variable, domaines formés d'un nombre *fini* d'intervalles). C'est cette conception nouvelle qui m'a permis d'exposer, d'une manière qui me paraît particulièrement simple, les déductions dont le principe se trouve dans mes premiers travaux sur la mesure et dans ma Note du 7 décembre 1903 (*voir* plus haut p. 239).

des polynomes. Mais il en est une qu'elle ne peut évidemment posséder : l'intégrale $f(x)$ ne peut définir la fonction intégrée $g(x)$ qu'avec une précision limitée; car deux fonctions, qui ne diffèrent qu'aux points d'un ensemble de mesure nulle, ont la même intégrale (leur différence est asymptotiquemeut équivalente à une suite de polynomes, dont tous les termes sont identiquement nuls).

On peut se proposer de rechercher la plus simple des fonctions qui admet comme intégrale $f(x)$; cet énoncé a-t-il un sens?

Si $g(x)$ a pour intégrale $f(x)$, on a, quels que soient x_1 et x_2,

$$f(x_2) - f(x_1) = \int_{x_1}^{x_2} g(x)\,dx.$$

La fonction $g(x)$, la plus simple qui satisfasse à cette condition, est évidemment la fonction constante

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Je n'insisterai pas sur l'étude des fonctions dérivées à laquelle on est ainsi conduit, car je n'aurais rien à ajouter d'essentiel aux beaux résultats de M. Lebesgue, dont le principal me paraît être que la dérivée de $f(x)$ existe, sauf en un ensemble de points de mesure nulle au plus; il semble donc naturel de choisir pour $g(x)$ cette dérivée aux points où elle existe; sa définition aux autres points (pourvu qu'elle reste bornée) ne modifie pas la valeur de l'intégrale (1).

La propriété fondamentale de l'intégrale est de définir ce qu'on peut appeler la valeur moyenne d'une fonction dans un domaine; il faut, pour cela, diviser la valeur de l'intégrale définie par la mesure du domaine. Cette valeur moyenne, lorsqu'on la connaît seule et que le domaine est suffisamment petit, doit être considérée comme la valeur probable de la fonction en un point quelconque, si l'on n'a sur cette fonction aucun autre renseignement. On voit que la considération des intégrales définies prouve que, si les fonctions tant soit peu compliquées ne sont pas calculables, du moins leurs valeurs probables dans un intervalle si petit qu'il soit, mais fini, existent toujours et sont calculables. C'est la raison pour laquelle on peut se borner à considérer les fonctions élémentaires calculables, tant qu'on n'a pas à envisager l'influence spéciale de points de discontinuité où la fonction devient infinie. C'est ce qu'on a fait généralement jusqu'ici dans

---

(1) Je signale deux Notes fort intéressantes de M. Denjoy, où ce jeune savant indique comment on peut résoudre d'une manière complète le problème inverse de la dérivation (intégration de toutes les fonctions dérivées) (*C. R. Acad. Sc.*, 1er et 22 avril 1912). Je me permets de signaler aussi la définition de la *dérivée en moyenne*, à laquelle j'ai été conduit par l'étude des fonctions qu'on rencontre dans la Mécanique statistique (*C. R. Acad. Sc.*, 29 avril 1912).

les applications des mathématiques à la philosophie naturelle; même dans les théories atomiques où interviennent des discontinuités, ce sont les valeurs moyennes qui jouent un rôle essentiel. Mais il n'est pas certain qu'on ne sera pas conduit à faire intervenir plus explicitement la discontinuité dans l'étude des phénomènes; c'est ce qui justifie l'étude, que nous allons faire, du cas des fonctions non bornées.

IV. — *L'intégration des fonctions non bornées.*

Une fonction non bornée peut admettre des points d'infinitude (où sa valeur est $+\infty$ ou $-\infty$) et des points d'indétermination; si l'ensemble de ces divers points n'était pas de mesure nulle, il est clair qu'on ne pourrait pas calculer l'intégrale de la fonction. Nous supposerons donc cet ensemble de mesure nulle. La fonction non bornée $f(x)$ est donc égale asymptotiquement à la limite de fonctions bornées; il suffit de considérer une fonction f_n égale à $f(x)$ si $|f(x)| < n$ et à $\pm n$ si $|f(x)| > n$, ou si $f(x)$ est infini ou indéterminé. Les fonctions $f_n(x)$ ont pour limite $f(x)$, sauf aux points d'infinitude ou d'indétermination; chacune de ces fonctions est bornée et bien définie, si $f(x)$ est défini analytiquement; on peut donc remplacer $f_n(x)$ par un polynome $P_n(x)$, tel que l'ensemble des $P_n(x)$ soit asymptotiquement équivalent à l'ensemble des $f_n(x)$, c'est-à-dire tende vers la limite $f(x)$, en excluant un ensemble de mesure aussi petite qu'on veut. On peut définir l'intégrale de $f(x)$ comme la limite de l'intégrale de $P_n(x)$, si cette limite existe; comme les P_n ne sont pas bornés, nous ne savons pas si cette limite existe. Cette définition est équivalente à la définition de M. Lebesgue.

Nous allons en donner une autre, qui donnera le même résultat quand elles s'appliqueront toutes deux, mais dont le champ d'application est plus étendu. Nous éviterons de considérer les séries divergentes (à somme infinie ou indéterminée); car on ne peut pas limiter l'étendue du domaine de divergence d'une telle série; par exemple, les séries

$$x + x + x + x + x + \ldots,$$
$$x - x + x - x + x - \ldots$$

divergent partout, sauf pour $x = 0$. Nous préférons, pour ce motif, considérer plutôt des séries dont les termes sont eux-mêmes des fonctions non bornées, ces séries étant généralement convergentes.

On sait comment on définit, dans les éléments, l'intégrale d'une fonction non bornée, telle que $f(x) = \dfrac{1}{\sqrt{x-c}}$; ou bien

$$f(x) = \frac{1}{x-c} \sin \frac{1}{x-c};$$

on pose :

$$\int_a^b f(x)\,dx = \lim_{\varepsilon=0}\int_a^{c-\varepsilon} f(x)\,dx + \lim_{\varepsilon'=0}\int_{c+\varepsilon'}^b f(x)\,dx.$$

L'intégrale existe, par définition, lorsque les limites existent. J'ai proposé de remplacer cette définition par la suivante; on sait que si l'on pose

$$h_i = x_i - x_{i-1}, \qquad x_{i-1} \leqq \xi_i \leqq x_i, \qquad x_0 = a, \qquad x_n = b,$$

on a, par définition, n augmentant indéfiniment et les h_i tendant vers zéro,

$$\int_a^b f(x)\,dx = \lim \sum h_i f(\xi_i).$$

Lorsqu'il y a un point singulier c, on entourera ce point d'un *intervalle d'exclusion*, $c-\varepsilon$, $c+\varepsilon'$, et l'on supposera que les points de division x_i sont tous extérieurs à cet intervalle. Si l'on a

$$x_{i-1} < c-\varepsilon < c+\varepsilon' < x_i,$$

on posera

$$h_i = x_i - (c+\varepsilon') + c - \varepsilon - x_{i-1} = x_i - x_{i-1} - \varepsilon - \varepsilon',$$

et l'on supposera ξ_i compris entre x_{i-1} et $c-\varepsilon$, ou entre $c+\varepsilon'$ et x_i, mais non entre $c-\varepsilon$ et $c+\varepsilon'$.

Si la *somme de Riemann*, ainsi définie, tend vers une limite, quels que soient ε et ε' et si cette limite tend elle-même vers une limite lorsque ε et ε' tendent vers zéro, cette dernière limite est, par définition, l'intégrale.

Il est manifeste que cette modification de la définition est purement formelle, lorsqu'il n'y a qu'un nombre limité de points singuliers; il n'en est pas de même, si les points singuliers sont denses dans un intervalle; il faut d'ailleurs supposer en ce cas, non seulement que chaque intervalle d'exclusion tend vers zéro, mais que leur somme tend aussi vers zéro. On est ainsi conduit à la définition suivante (qui s'étendrait sans changement au cas des intégrales multiples, mais qui, en ce cas, paraît moins utile) :

Soit $f(x)$ une fonction non bornée, non intégrable au sens de Riemann; supposons qu'on puisse déterminer dans le champ d'intégration une infinité énumérable de points $A_1, A_2, \ldots, A_n, \ldots$ ayant la propriété suivante : si l'on entoure le point A_n d'un intervalle d'exclusion B_nC_n, tel que la série ΣB_nC_n soit convergente et de somme ε, les sommes de Riemann généralisées tendent vers une limite, quels que soient les intervalles, et cette limite tend elle-même vers une limite, lorsque ε tend vers zéro; cette dernière limite est, par définition, l'intégrale, au sens de Riemann, géné-

ralisée. Les sommes de Riemann généralisées sont les sommes

$$\Sigma h_i f(\xi_i),$$

dans lesquelles on suppose :

1° Que les points de division x_i n'appartiennent pas aux intervalles d'exclusion;

2° Que h_i est égal à $x_i - x_{i-1}$, diminué, s'il y a lieu, de la longueur des intervalles d'exclusion;

3° Que ξ_i est situé entre x_{i-1} et x_i, mais n'appartient pas non plus aux intervalles d'exclusion.

On étudie la limite de ces sommes de Riemann dans l'hypothèse que les h_i tendent vers zéro et l'on fait tendre ensuite vers zéro les intervalles d'exclusion.

V. — *Comparaison avec l'intégrale de M. Lebesgue.*

Il est aisé de voir que, dans le cas des intégrales multiples, la nouvelle définition est équivalente à celle de M. Lebesgue; on sait, en effet, que l'intégrale multiple d'une fonction ne peut converger que si l'intégrale de la valeur absolue converge; et, dans une somme absolument convergente, l'ordre des termes est indifférent.

Au contraire, dans le cas des intégrales simples, l'intégrale, au sens de M. Lebesgue, ne peut exister que si l'intégrale de la valeur absolue existe, tandis qu'il n'en est pas de même pour l'intégrale au sens de Riemann. Ceci tient, on le sait, à ce que la définition de Riemann conduit à une série dont la convergence peut être assurée par l'alternance des signes. Tel est le cas pour la fonction $\frac{\sin x}{x}$ à l'infini ou pour la fonction $\frac{1}{x} \sin \frac{1}{x}$ pour $x = 0$.

Il est clair que la propriété d'être intégrable au sens de M. Lebesgue, c'est-à-dire d'avoir une valeur absolue intégrable, étant une propriété plus restrictive, est, par cela même, plus commode à utiliser dans certaines applications, de même que les séries absolument convergentes sont plus aisées à manier que les séries simplement convergentes. Mais ce n'est peut-être pas une raison pour qu'on néglige complètement l'étude des intégrales ou des séries qui ne convergent pas absolument.

Si deux fonctions sont intégrables au sens de Riemann généralisé, et si leur somme est intégrable, l'intégrale de la somme est égale à la somme des intégrales. Mais la somme peut ne pas être intégrable avec notre définition, parce que les intervalles d'exclusion de l'une des fonctions peuvent être, dans des cas particuliers, choisis précisément de manière à rendre divergente l'intégrale de l'autre fonction. On peut convenir d'étendre la définition et de dire que : si une fonction n'est pas intégrable d'après la définition du paragraphe précédent, mais peut être décomposée

en la somme de deux fonctions intégrables d'après cette définition, son intégrale est, par définition, la somme des intégrales de ces fonctions. Pour que cette nouvelle extension de la notion d'intégrale de Riemann généralisée ne soit pas contradictoire, il faut faire voir qu'elle ne peut pas conduire à attribuer deux valeurs différentes à l'intégrale d'une même fonction. Or, c'est ce qui est aisé de démontrer. Soit

$$f = f_1 + f_2 = g_1 + g_2;$$

par hypothèse, les intégrales riemanniennes généralisées de f_1 et de f_2 existent, et aussi celles de g_1 et de g_2. Il s'agit de prouver qu'on a

$$\int f_1\,dx + \int f_2\,dx = \int g_1\,dx + \int g_2\,dx.$$

Désignons respectivement par (A_1), (A_2), (B_1), (B_2) les ensembles de points singuliers qui interviennent dans les définitions des intégrales de f_1, f_2, g_1, g_2. Par définition, à tout nombre ε correspondent des nombres ε_1, ε_2, η_1, η_2, tels que si l'étendue des intervalles d'exclusion correspondant aux (A_1) est inférieure à ε_1, la limite des sommes riemanniennes correspondantes diffère de moins de ε de l'intégrale $\int f_1\,dx$: mais, si l'on excluait, dans le calcul de cette intégrale, des intervalles arbitraires correspondant aux (A_2), (B_1), (B_2), on ne sait pas si la convergence subsisterait. Ce qu'il suffit de faire voir, c'est qu'il est possible de changer infiniment peu la valeur de cette intégrale par un *choix particulier* des intervalles d'exclusion (B_1), ce choix particulier étant assujetti seulement à avoir une étendue inférieure à η_1. Or, cela est évident, car l'intégrale riemannienne généralisée d'une fonction f_1 possède évidemment la propriété fondamentale

$$\int_a^b f_1\,dx = \int_a^c f_1\,dx + \int_c^b f_1\,dx,$$

chacune des intégrales du second membre existant si celle du premier membre existe. Les intégrales de f_1 dans l'intérieur de chacun des intervalles d'exclusion (B_1) existent donc, et l'on peut prendre chacun de ces intervalles assez petit pour que la valeur correspondante de l'intégrale soit aussi petite qu'on veut.

Il est donc possible de faire un *choix particulier* des intervalles d'exclusion (A_1), (A_2), (B_1), (B_2), tels que chacune des intégrales soit calculée, à moins de 2ε près, lorsqu'on exclut tous ces intervalles; il en résulte que

$$\int f_1\,dx + \int f_2\,dx - \int g_1\,dx - \int g_2\,dx$$

diffère de moins de 8ε de la même expression, où les intégrales seraient étendues seulement au domaine qui subsiste lorsqu'on enlève tous les

intervalles d'exclusion; or, cette expression est alors égale à

$$\int (f_1+f_2-g_1-g_2)\,dx = 0.$$

L'égalité à démontrer est donc vérifiée à 8ε près; elle est donc rigoureusement établie.

Comme exemple de fonctions f_1 et f_2, telles que leur somme ne soit pas directement intégrable alors que chacune d'elles l'est, on peut prendre

$$f_1(x) = \frac{1}{x}\sin\frac{\pi}{x},$$

$$f_2(x) = \sum_1^{\infty} \frac{e^{-n}}{\sqrt{2-(2n+1)x}}.$$

Si l'on désirait considérer une infinité de fonctions non bornées telles que, non seulement la somme, mais les produits deux à deux de deux fonctions quelconques de l'ensemble appartiennent à l'ensemble, il faudrait partir d'un nombre fini (ou infini) de fonctions telles que chacune soit intégrable ainsi que toutes ses puissances. Tel est le cas des fonctions telles que la suivante :

$$\sum e^{-n}\log\left|x-\frac{1}{n}\right|.$$

On est naturellement conduit par ce qui précède à étudier les séries généralement convergentes de fonctions non bornées, c'est-à-dire les séries qui convergent presque partout, les portions exclues étant les points singuliers et un certain entourage asymptotique de ces points. C'est en réalité cette étude qui m'avait conduit à considérer, pour la première fois, des ensembles de mesure nulle et à utiliser leurs principales propriétés, étude d'où j'ai déduit ensuite la définition précise de la mesure. Observons à ce sujet que, si l'on suppose que l'ensemble des points de divergence soit de mesure nulle, le troisième théorème fondamental subsiste, avec cette différence toutefois que, lorsque α tend vers o, les polynomes P_n ne sont pas bornés dans leur ensemble.

On n'est donc pas assuré de l'existence de l'intégrale; il faut effectivement que la fonction ne croisse pas trop vite dans le voisinage des points singuliers. La convergence est assurée, si l'ordre des points singuliers est inférieur à un nombre fixe inférieur à 1; elle peut être réalisée aussi dans certains cas où cet ordre est asymptotique à l'unité (par valeurs inférieures), par exemple, pour la fonction

$$\sum_{n=1}^{\infty} \frac{e^{-n}}{\left(x-\frac{1}{n}\right)^{1-\frac{1}{n}}}.$$

CHAPITRE III.

Le calcul effectif des intégrales.

I. — *Le calcul par la définition.*

Pour qu'une intégrale puisse être effectivement calculée par les méthodes précédentes, il est évidemment nécessaire et suffisant que l'ensemble des points où la fonction à intégrer n'est pas calculable soit de mesure nulle. Ceci pourrait sembler impossible, si l'on considère qu'une fonction ne peut être calculable que pour les valeurs calculables des variables et que l'ensemble de ces valeurs est de mesure nulle; mais l'on peut regarder une fonction comme calculable pour certaines valeurs non calculables des variables, si l'on sait que, pour ces valeurs, la fonction coïncide avec une fonction continue calculable. A ce point de vue, notre troisième théorème fondamental apparaît comme indispensable; sans ce théorème (ou une proposition équivalente), il serait complètement dénué de sens de parler d'intégrale d'une fonction discontinue, car les opérations par lesquelles pourrait être calculée cette intégrale seraient absolument inexécutables, non seulement au point de vue pratique, mais au point de vue théorique : je veux dire qu'on ne pourrait imaginer aucun moyen, si long fût-il, de les exécuter.

Voici, en nous bornant à une variable pour simplifier l'écriture, comment se pose pratiquement le problème. Une fonction $f(x)$ est définie comme la limite de fonctions connues (ou du moins plus simples) $f_1(x), f_2(x), \ldots, f_n(x), \ldots$. Nous admettons qu'on sache calculer les intégrales de chacune de ces fonctions avec une approximation donnée. Le problème de l'intégration de $f(x)$ est immédiatement résolu dans le cas où la suite $f_n(x)$ converge uniformément et où l'on a la mesure de cette convergence uniforme. S'il n'en est pas ainsi, on devra tout d'abord chercher à obtenir une limite supérieure M du module de f et des f_n. On cherchera ensuite, au moyen de la démonstration même du fait que f_n tend vers une limite, à déterminer un nombre ε_n supérieur *en général* à $f - f_n$, et l'on désignera par α_n l'étendue du domaine où $f - f_n$ dépasse ε_n; on aura alors

$$\left| \int_a^b f\,dx - \int_a^b f_n\,dx \right| < \alpha_n \mathrm{M} + \varepsilon_n (b - a).$$

On aura une approximation déterminée en prenant n assez grand pour que ε_n et α_n soient inférieurs à des nombres fixés, et en calculant l'intégrale de f_n avec une approximation suffisante.

On voit que le calcul sera le plus avantageux possible, si l'on sait s'arranger pour que ε_n et α_n soient sensiblement du même ordre de grandeur. Il est difficile de donner sur ce point des indications générales. Mais les

choses se simplifient beaucoup si l'on se borne à la considération d'une catégorie déterminée de fonctions, même si cette catégorie est très étendue. On pourrait aisément reprendre à ce point de vue l'exposition donnée dans cette Note (*voir* plus haut Chap. I, § II).

II. — *L'emploi des probabilités dénombrables.*

Pour calculer certaines intégrales, il est plus commode d'utiliser la notion de probabilité. J'ai montré ailleurs [1] comment on peut évaluer la probabilité pour qu'un nombre x (compris entre 0 et 1) satisfasse à une condition déterminée, ce nombre x étant défini par une infinité dénombrable de conditions simples, dont les probabilités respectives sont connues. Cette théorie se superpose tout à fait à la théorie de la mesure; de même, on pourrait développer une théorie de l'espérance mathématique dans les cas dénombrables qui se superposerait à la théorie de l'intégrale définie. Le premier cas revient à ne s'occuper que de l'intégration des fonctions prenant seulement les valeurs 0 et 1; le problème de l'intégration de ces fonctions est identique au problème de la mesure des ensembles. J'ai indiqué des exemples de calculs de ce genre dans la Note que je viens de citer. Je n'y reviendrai pas; je voudrais seulement signaler brièvement une question qui me paraît intéressante : est-il possible de définir une fonction $f(x)$ qui soit intégrable au moyen des probabilités dénombrables, sans pouvoir être ramenée aux fonctions calculables? Tel serait le cas pour une fonction telle que, dans tout intervalle si petit qu'il soit, les valeurs 0 et 1 seraient également probables. Son intégrale entre les limites 0 et 1 serait alors évidemment égale à $\frac{1}{2}$, tandis que sa valeur, si l'on n'en savait rien de plus que ce que nous venons de dire, ne pourrait être connue pour aucune valeur de la variable (et, en tous cas, quelle que soit sa définition, ne pourrait jamais être connue pour toutes les valeurs, mais seulement au plus pour une infinité dénombrable de valeurs de la variable).

Mais il ne semble pas qu'il soit possible d'arriver, au moyen des probabilités dénombrables, à la définition d'une telle fonction; on se trouve, en effet, soit dans le cas de convergence, soit dans le cas de divergence, et la probabilité limite (désignée par A_∞ dans la Note citée) est égale, suivant le cas, à 0 ou à l'unité, sans avoir jamais une valeur intermédiaire. Il faudrait donc tout au moins compliquer notablement les définitions de cette Note et arriver à définir des probabilités successives dépendant les unes des autres suivant une loi suffisamment compliquée pour que l'on soit à la fois dans le cas de convergence et dans le cas de divergence. Mais cela ne paraît pas aisé.

Ce résultat peut paraître contradictoire avec celui qu'a obtenu M. Lebesgue, qui est arrivé à « nommer » une fonction non définissable

---

[1] Note V de cet Ouvrage. — *Voir* aussi *Comptes rendus de l'Académie des Sciences*, t. CLIV, 29 avril 1912, p. 1150.

analytiquement; mais il faut observer que la fonction de M. Lebesgue ne diffère qu'aux points d'un ensemble de mesure nulle d'une fonction définissable analytiquement; elle est donc équivalente à une telle fonction au point de vue de l'intégration [1].

La question peut être posée sous une forme équivalente, peut-être plus saisissante : est-il possible, ou non de définir avec propriété *asymptotique* des nombres irrationnels, tels que cette propriété et son contraire soient également probables ? Par propriété asymptotique on entend, s'il s'agit, par exemple, de nombres décimaux, une propriété dont la définition ne dépend pas des n premières décimales, quel que soit n, c'est-à-dire reste la même pour tous les nombres dans lesquels ces n premières décimales seules diffèrent. On peut dire aussi qu'une propriété asymptotique est entièrement homogène au continu, c'est-à-dire se présente sous la même forme pour deux intervalles égaux se déduisant l'un de l'autre par une translation commensurable. Il me paraît résulter, des considérations développées dans ce Mémoire et de la théorie des probabilités dénombrables, que les seuls ensembles homogènes au continu sont les ensembles de mesure nulle; on ne peut obtenir des ensembles de mesure non nulle qu'en faisant intervenir explicitement certains intervalles, c'est-à-dire que les seuls ensembles de mesure non nulle qui puissent être définis sont mesurables.

Je ne me dissimule pas que les raisons que j'invoque à l'appui de cette conclusion ne paraîtront pas satisfaisantes à tous; il me semble néanmoins qu'elle s'impose à ceux qui considèrent qu'une définition mathématique doit fournir quelque procédé de calcul.

COMPLÉMENT AUX NOTES V ET VI.

Pendant l'impression de cet Ouvrage, M. Hausdorff a publié dans les *Mathematische Annalen* (t. LXXV, p. 428) un résultat qui me paraît confirmer diverses idées émises dans les Notes V et VI. Voici comment on peut exposer ce résultat.

Soit une sphère S, de surface égale à 1; désignons par φ une rotation de π autour d'un diamètre et par ψ une rotation de $\frac{2}{3}\pi$ autour d'un autre diamètre faisant un angle θ avec le premier. Considérons toutes

---

[1] Je laisse ici de côté les objections qu'on peut faire à l'*existence* de la fonction « nommée » par M. Lebesgue, qui a d'ailleurs insisté lui-même sur la distinction entre *nommer* et *définir* une fonction; je serais volontiers plus catégorique que lui : partout où devraient intervenir effectivement *tous* les nombres transfinis de seconde classe (et non pas seulement ceux qui sont inférieurs à l'un d'eux fixé d'avance), on me paraît sortir du domaine des Mathématiques. [*Voir* mon article sur *La philosophie des Mathématiques et l'infini* (*Revue du mois*, 10 août 1912). — Ci-dessus, Note IV, § VII.]

[1] *Sur les relations entre ces considérations et certaines théories de Physique mathématique*, *voir* ma Note : *Sur les probabilités et les hypothèses de discontinuité* (*C. R. Acad. Sc.*, 5 janvier 1914).

les rotations R de la forme $\psi^{m_0}\varphi\psi^{m_1}\varphi\psi^{m_2}\varphi\ldots\psi^{m_{n-1}}\varphi\psi^{m_n}$, les nombres $m_1, m_2, \ldots, m_{n-1}$ étant 1 ou 2, m_0 et m_n étant 0, 1 ou 2. Ces rotations sont énumérables; on peut choisir θ de manière qu'elles soient toutes distinctes, car l'ensemble des valeurs de θ vérifiant les équations obtenues en les égalant deux à deux est énumérable. L'angle θ étant ainsi choisi, chacune de ces rotations laisse invariants deux points seulement; soient Q l'ensemble (énumérable) de ces points, P l'ensemble des autres points de la sphère; la mesure de P est 1. Soient x un point de P, X l'ensemble des points qui se déduisent de x par les rotations R; l'ensemble X est énumérable. Les points de P sont ainsi répartis en une infinité (non énumérable) d'ensembles tels que X; DANS CHACUN DE CES ENSEMBLES X, CHOISISSONS UN POINT x; l'ensemble des points ainsi choisis forme un ensemble M. Nous allons répartir tous les points de P entre trois ensembles A, B, C :

Considérons [1] les produits de M par les rotations R considérées plus haut, l'entier n étant positif ou nul; un tel produit sera rangé dans A si $m_n = 0$, dans B si $m_n = 1$ et dans C si $m_n = 2$, *avec toutefois l'exception suivante :* les produits de la forme $M(\varphi\psi^2)^k$, k étant un entier positif ou nul, seront rangés dans A, les produits $M(\varphi\psi^2)^k\varphi$, dans B et les produits $M(\varphi\psi^2)^k\varphi\psi$ dans C. On voit aisément que tout point de P se trouve classé sans ambiguïté dans l'un des ensembles A, B, C et que l'on a

$$P = A + B + C, \qquad A\psi = B, \qquad A\psi^2 = C, \qquad A\varphi = B + C.$$

Si donc on désigne par a, b, c les probabilités pour qu'un point de S appartienne à A, B, C et si l'on admet que la probabilité pour qu'un point appartenant à un ensemble E ne change pas par une rotation autour d'un diamètre (c'est ce que M. Lebesgue exprime en disant que deux ensembles congruents ont même mesure), on a les égalités contradictoires :

$$a + b + c = 1, \qquad a = b, \qquad a = c, \qquad a = b + c.$$

La contradiction a son origine dans l'application (écrite en capitales plus haut) de l'*axiome du choix* de M. Zermelo. L'ensemble A est homogène sur la sphère; mais il en est à la fois la moitié et le tiers; on recouvre la surface S (à l'exclusion de l'ensemble énumérable Q) avec l'ensemble A et ceux que l'on en déduit, soit par la seule rotation φ, soit par les deux rotations ψ et ψ^2. Le paradoxe vient de ce que A *n'est pas défini*, au sens logique et précis du mot *défini*. Si l'on fait fi de la précision et de la logique, on est conduit à des contradictions.

---

(1) Je modifie ici l'exposition de M. Hausdorff afin de donner explicitement les ensembles A, B, C. Cela permet de généraliser, en prenant la rotation ψ incommensurable avec π. Les entiers $m_0, m_1, m_2, \ldots, m_n$ sont alors positifs ou négatifs (m_0 et m_n pouvant être nuls); on classe MR dans A_k si $m_n = k$, avec l'exception suivante : les ensembles $M(\varphi\psi)^n\varphi\psi^{1+k}$ sont dans A_k et non dans A_{1+k}.

On a alors $A_0\varphi = \sum_{\pm 1}^{\pm\infty} A_k$; $A_k\psi = A_{k+1}$.

NOTE VII.

POUR ET CONTRE LA LOGIQUE EMPIRIQUE (¹).

I. — Logique formelle et logique empiriste (²).

Dans un article précédent je me suis efforcé de montrer comment le débat d'aujourd'hui sur l'emploi du principe du tiers exclu peut se rattacher à d'autres débats plus anciens et pourquoi il en constitue, en somme, l'issue fatale. Faisons abstraction du côté historique de la question pour insister sur son aspect logique. Pénétrons dans le vif du sujet en traitant un exemple précis.

Disons qu'un nombre x est algébrique, s'il est possible de déterminer une équation algébrique à coefficients entiers

$$a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n = 0$$

dont il soit racine.

Disons que le nombre x est transcendant si l'hypothèse de son algébricité peut être réduite à l'absurde.

Ceci étant, le nombre $r = x^2$ est-il nécessairement soit algébrique, soit transcendant ? Cette alternative n'est-elle pas sujette à caution ? Elle n'est pas tranchée, que je sache; est-elle bien légitime ?

Il n'est pas exclu que l'on ne puisse jamais démontrer que le nombre r est algébrique ou bien encore qu'il est transcendant. Il faut envisager la circonstance suivante :

Éventualité E : Il n'existe aucune démonstration permettant d'établir la transcendance ou bien encore l'algébricité du nombre r.

Cette éventualité ne peut être écartée d'un revers de main et la proposition — le nombre r est algébrique — pourrait n'être ni démontrable ni réductible à l'absurde.

(¹) Il m'a paru intéressant de réunir ici des discussions auxquelles ont donné lieu récemment les théories de M. Brouwer sur la logique empiriste. Je remercie MM. R. Wavre, Paul Lévy, M. Barzin et A. Errera de m'avoir autorisé à reproduire leurs intéressants articles sur ce sujet.

(²) *Revue de Métaphysique et de Morale*, janvier 1926.

Devant un problème irrésolu on a bien le droit, je pense, de se demander si le problème est résoluble; et à moins d'admettre par principe que tout problème mathématique bien posé est résoluble, il faut envisager l'insolubilité de certaines questions. Ainsi le problème que pose le nombre r n'étant pas résolu, on est en droit de se demander si l'éventualité E n'est pas réalisée.

Il suffit donc de fournir l'exemple d'un nombre dont on ne sache s'il est algébrique ou transcendant pour fournir en même temps l'exemple d'un nombre qui, jusqu'à plus ample information, *pourrait* n'être ni l'un ni l'autre. Mais, d'autre part, il serait vain, me semble-t-il, de vouloir définir un nombre qui ne *soit* ni algébrique ni transcendant, car la seule manière de prouver qu'il n'est pas algébrique consistant à prouver qu'il serait absurde qu'il le fût, ce nombre serait transcendant.

En d'autres termes, il me semble impossible de définir un nombre donnant lieu à l'éventualité E, mais il n'est pas exclu que cette éventualité se présente sans que nous le sachions. Il est aussi chimérique de vouloir la réaliser pour un nombre particulier que de vouloir l'écarter *a priori* pour tout nombre. Et la question — l'éventualité E se présente-t-elle ? — est de celles que les mathématiciens renvoient aux philosophes. Et, cependant, le mathématicien qui veut affirmer qu'un nombre, quel qu'il soit, est toujours algébrique ou transcendant, doit en avoir le cœur net; il doit prendre position.

Ici se manifeste une divergence de vue, les deux personnages dont parlait Paul Du Bois Reymond rentrent en scène.

Celui qui ne tient pour vrai que ce qui est effectivement démontrable, et pour absurde que ce qui se laisse effectivement réduire à une contradiction — l'empiriste, en un mot — pourra douter que tout nombre doive être algébrique ou transcendant, et il pourra fort bien proclamer : la proposition — le nombre r est algébrique — pourrait n'être ni vraie ni absurde.

Celui qui identifie le vrai avec le non contradictoire — l'idéaliste ou le formaliste, ce dernier terme convient mieux ici, — proclamera algébrique, tous les nombres dont on est assuré de ne pouvoir démontrer la transcendance et, en effet, une proposition dont on ne peut par hypothèse démontrer l'absurdité peut toujours être formellement tenue pour vraie.

Enfin le formaliste qui, sans aller jusque-là, dit, avec Poincaré : « Le mot *exister* ne peut avoir qu'un sens, il signifie exempt de contradiction » proclamera algébrique tout nombre prouvé non transcendant.

Ce faisant, le formaliste n'hésite pas à proclamer un nombre algébrique, ou à affirmer l'existence d'une équation, sans être assuré de pouvoir trouver l'équation elle-même dont le nombre serait racine. Ce nombre ne serait que formellement, idéalement algébrique; il ne le serait qu'en transformant tant soit peu le sens de la définition du début. Et, en effet, le formaliste pose cette définition d'une autre manière, il dit :

« Un nombre est algébrique s'il existe une équation algébrique à coefficients entiers dont il soit racine, transcendant s'il n'en existe aucune. »

Mais cette définition formaliste cache sous son aspect formel toute la difficulté, puisque la seule manière de mettre tout le monde d'accord consiste-

rait à prouver qu'il existe une équation, en la trouvant, et qu'il n'en existe aucune, en démontrant qu'il serait effectivement absurde qu'il en existât une. L'empiriste exige, en effet, que l'on explicite le sens concret des propositions purement existentielles, telles que « il existe » ou « il n'existe pas ». C'est ce point que nous avons approfondi dans l'article rappelé.

La seconde définition n'est qu'un maquillage; son aspect formel laisse croire que l'alternative s'impose *a priori*, et nous venons de voir que, sur le plan mathématique, elle pourrait, dans l'éventualité E, ne pas s'imposer. L'un des personnages admet la légitimité de l'alternative formelle purement existentielle « il existe » ou « il n'existe pas » et applique à la lettre les principes de la logique classique; l'autre cherche à pénétrer le sens concret de l'alternative et il estime que, pour être plus logique encore, il est prudent de ne pas se laisser illusionner par le caractère formel des principes de la logique; il ne les applique qu'à bon escient. Nous sommes sur le chemin d'un aphorisme en disant qu'il ne faut employer la logique formelle qu'avec précaution. Eh bien! Tout objet serait-il éclairé, parce que nous projetons deux lumières distinctes ?

Une alternative n'est légitime que si l'on est assuré de pouvoir la trancher dans tous les cas particuliers. Or, la trancher signifie prouver directement la vérité de l'une de ses parties et non prouver l'absurdité de l'une pour conclure à la vérité de l'autre, ce qui présupposerait l'alternative légitime. Mais si l'alternative n'est légitime qu'une fois tranchée par une démonstration où elle n'entre pas, alors pratiquement elle est inutile et l'on ne retiendra qu'une de ses parties. Ainsi l'emploi du mot « ou » dans une proposition générale n'est licite que si l'on est assuré de pouvoir s'en passer dans tous les cas particuliers. L'empirisme implique donc un maniement très précautionneux des propositions générales.

(C'est ainsi qu'il faudrait douter, par exemple, du principe de Bolzano-Weierstrass : — Si une infinité de points sont répartis sur un intervalle fermé de longueur finie, il existe toujours un point-limite sur l'intervalle, — principe qui est fondamental dans l'analyse. L'empirisme, en effet, prescrit la construction d'un tel point-limite. Le principe de Bolzano-Weierstrass, fondé par ailleurs sur l'alternative : fini ou infini, n'en serait pas moins une incitation à extraire pratiquement, dans chaque cas particulier, de l'infinité des points, une suite convergente. Et alors, dans chacun des cas où le principe de cette extraction aura pu être formulé explicitement, l'existence du point-limite sera établie.

L'intransigeance de l'empiriste s'accentue et, par exemple, s'il doutait, il y a 20 ans, que le continu pût être *bien ordonné*, il doute aujourd'hui qu'il puisse être tout simplement *ordonné*.

Faut-il aller plus loin et voir dans cette attitude empiriste une tentative d'affranchissement à l'égard des règles formelles de la logique ? Le principe du tiers exclu aurait-il un caractère trop scolastique ?

Qui voudrait tenter de s'affranchir des principes de la logique formelle

devrait employer une langue toute spéciale, nos langues modernes sont toutes imprégnées de logique traditionnelle.

La mathématique serait-elle cette langue, pourrait-elle briser ce cadre ? En fait, l'intuition ne s'en embarrasse guère, et les principes de la logique formelle ne sont pas d'un très grand secours dans l'invention mathématique.

Ce n'est pas que ces principes soient, comme on le dit quelquefois, les conditions de toute pensée viable; ils sont plutôt des principes de délimitation précise, mais toute formelle des pensées.

Quel est donc leur véritable rôle ?

Le principe de contradiction et celui du tiers exclu sont comme les ciseaux parfaitement aiguisés qui permettent de détacher formellement A de non A. Intimement liés à l'emploi de la négation, ils le réglementent et se placent, à ce titre, au premier rang des règles de syntaxe. Et quand on veut écrire, il vaut mieux les respecter. Ainsi on ne peut affirmer A et non A, il faut affirmer A ou non A. Cela est fort bon, mais l'emploi de la négative une fois réglementé, toute la science reste à faire. Et l'on peut se demander si la vérité démontrable et l'absurdité démontrable sont bien dans le rapport de A et de non A. Non, nous l'avons vu, il faut en faire quelque chose comme Pierre et Paul, ou mieux A et B, ces deux termes étant entre eux dans une relation logique toute spéciale.

Distinguons la logique formelle, logique de la négation ou encore du vrai et du faux, dans laquelle l'alternative s'impose entre le vrai et le faux, et *la logique mathématique empiriste du vrai et de l'absurde*, dans laquelle l'alternative entre le vrai et l'absurde ne s'imposerait plus; vrai signifiant l'effectivement démontrable et absurde, ce qui peut être effectivement réduit à une contradiction, la contradiction étant prise dans le sens formel. L'absurde implique le faux, mais le faux n'implique pas toujours l'absurde.

Ainsi, la *logique mathématique empiriste ne serait qu'une logique appliquée ayant ses principes propres.*

Mais peut-on songer à formuler des principes généraux de logique appliquée? Ne sont-ils pas aussi nombreux que les applications elles-mêmes ? Et ne sont-ils pas à la merci de quelques nouvelles difficultés mathématiques, de quelque subtilité qui contraindrait à révoquer les principes qui nous paraissent les plus irrévocables ? Ceci n'est donc qu'une suggestion. Rappelons quelques axiomes de la logique classique; formulons quelques principes de la logique du vrai et de l'absurde, dans le but de faire apparaître la différence. Adressons-nous au calcul des propositions, puisque c'est précisément la vérité et l'absurdité des propositions qui nous préoccupent.

LES PRINCIPES SEMBLABLES.

Les notions de proposition et d'implication sont prises comme notions premières. Soient A et B deux propositions. A implique B s'écrira A — B, et s'interprétera : si A est vraie, B est vraie.

Nous dirons que deux propositions sont équivalentes si elles s'impliquent mutuellement.

1. *Principe du syllogisme.* — Si une proposition en implique une autre, qui en implique à son tour une troisième, la première implique la troisième.

2. *Principe de la déduction.* — Si une proposition est vraie, et si elle en implique une autre, celle-ci peut être affirmée vraie isolément.

3. *Principe de la contradiction.* — *Énoncé formel :* Une proposition ne peut être vraie et fausse.

Énoncé empiriste : Une proposition ne peut être vraie et absurde.

4. *Principe de l'implication du faux et de l'implication de l'absurde.* — *Énoncé formel :* Ce qui implique le faux est faux.

Énoncé empiriste : Ce qui implique l'absurde est absurde.

LES PRINCIPES DISSEMBLABLES.

5. *Principe du tiers exclu.* — *Énoncé formel :* Une proposition ne peut être que vraie ou fausse.

— Le principe n'a pas d'équivalent en logique empiriste.

6. *Principe du prédicat de fausseté et du prédicat d'absurdité.* — *Énoncé formel :* La fausseté de la fausseté d'une proposition équivaut à la vérité de cette proposition.

Énoncé empiriste : La vérité d'une proposition implique l'absurdité de l'absurdité de cette proposition.

Pour bien marquer la différence, explicitons trois principes que l'on rejette en logique empiriste, et qui seraient traduits dans le langage du vrai et du faux, conséquences des énoncés formels 5 et 6.

α. Une proposition est vraie ou absurde.

β. L'absurdité de l'absurdité d'une proposition implique la vérité de cette proposition.

γ. Une proposition est absurde ou bien il est absurde qu'elle soit absurde.

M. Brouwer rejette bien les principes α, β et γ, mais il n'énumère pas les principes qu'il admet.

Cette tâche, d'ailleurs, ne lui incombe pas, et voici pourquoi :

Les principes de logique ne sont pour lui que des propriétés du langage mathématique. Ce langage est une activité matérielle amathématique, servant à communiquer ou à registrer quelque épisode de l'activité mentale averbale qu'est la pensée mathématique.

Il présente toutes les défectuosités des langages servant à communiquer ou à registrer un domaine quelconque de la pensée humaine.

Les principes de logique, propriétés du langage, constituent donc pour

M. Brouwer non un phénomène mathématique, mais un phénomène ethnographique.

Ce point de vue, s'il était partagé un jour par les mathématiciens, créerait une véritable révolution philosophique.

Pour ma part, je ne puis l'adopter intégralement, et ce n'est qu'en poussant l'empirisme du premier des personnages de Du Bois Reymond jusqu'à ses conséquences inévitables, que je puis douter, avec et après M. Brouwer, de la légitimité des énoncés α, β, γ.

Comparons quelques conséquences des principes posés plus haut.

POINT DE VUE FORMEL.

Si A présente la vérité d'une proposition, $\overline{A}$ représentera sa fausseté. Et l'on peut faire de $\overline{A}$ une nouvelle proposition qui sera la négative; ou simplement la négation de A, non A. Sous cet aspect, le principe 6 n'est autre que celui de la double négation. Deux négations équivalent à une affirmation $\overline{\overline{A}}$ équivaut à A. Ainsi l'adjonction répétée d'un trait ne peut fournir que les deux propositions A et $\overline{A}$.

Le principe 4 peut s'interpréter ainsi : Si A implique B, la fausseté de B implique la fausseté de A, et s'écrire :

$$(a) \qquad (A \rightarrow B) \rightarrow (\overline{B} \rightarrow \overline{A}).$$

Sous cette forme on pourrait l'appeler principe de contraposition. On en déduit :

$$(\overline{B} \rightarrow \overline{A}) \rightarrow (\overline{\overline{A}} \rightarrow \overline{\overline{B}}),$$

de sorte que

$$(b) \qquad (\overline{B} \rightarrow \overline{A}) \rightarrow (A \rightarrow B).$$

Cette dernière relation est le schème des démonstrations par l'absurde A étant supposée vraie, on établit que la fausseté de B entraînerait la fausseté de A, ce qu'il faut rejeter, et alors B, qui ne peut être fausse, est vraie.

Les relations (a) et (b) montrent que les deux implications : $A \rightarrow B$ et $\overline{B} \rightarrow \overline{A}$, sont équivalentes.

POINT DE VUE EMPIRISTE.

Si A représente la vérité d'une proposition, $\overset{1}{A}$ représentera son absurdité, $\overset{2}{A}$ l'absurdité de son absurdité ou absurdité seconde, $\overset{3}{A}$ l'absurdité de son absurdité seconde ou absurdité troisième.

Le principe 6 s'exprime $A \rightarrow \overset{2}{A}$, tandis que le principe 3, que nous rejetons, s'écrirait $\overset{2}{A} \rightarrow A$.

Le principe 4 peut s'interpréter ainsi : Si A implique B, l'absurdité de B implique l'absurdité de A et s'écrire :

$$(c) \qquad (A \rightarrow B) \rightarrow \left(\overset{1}{B} \rightarrow \overset{1}{A}\right).$$

On pourrait encore l'appeler : principe de contraposition.

Or, de la relation $A \rightarrow \overset{2}{A}$, principe 6, on déduit par contraposition :

$$\left(A \rightarrow \overset{2}{A}\right) \rightarrow \left(\overset{3}{A} \rightarrow \overset{1}{A}\right),$$

et l'implication $\overset{3}{A} \rightarrow \overset{1}{A}$ peut être posée vraie isolément en vertu du principe de la déduction. Mais, d'autre part le principe 6 appliqué à la proposition $\overset{1}{A}$, s'écrit $\overset{1}{A} \rightarrow \overset{3}{A}$. L'absurdité troisième équivaut à l'absurdité première. Ce résultat fut exposé en 1923 par M. Brouwer. Il n'y a donc que deux propositions que l'on puisse obtenir par répétition du prédicat d'absurdité, ce sont $\overset{1}{A}$ et $\overset{2}{A}$, avec A cela fait trois propositions distinctes. Remarquons en passant que rien ne nous assure que l'une des trois A, $\overset{1}{A}$ ou $\overset{2}{A}$, doive être vraie.

Et voici une conséquence immédiate de ce qui précède.

La démonstration par l'absurde n'est plus valable en général; il ne faut pas s'en étonner puisqu'elle consisterait à inférer la vérité de l'absurdité seconde, inférence β, qui est rejetée. Examinons-la tout de même; je vais faire voir qu'elle reste valable dans un cas particulier.

Le principe de contraposition permet d'écrire :

$$\left(\overset{1}{B} \rightarrow \overset{1}{A}\right) \rightarrow \left(\overset{2}{A} \rightarrow \overset{2}{B}\right);$$

alors supposons que l'implication $\overset{1}{B} \rightarrow \overset{1}{A}$ soit établie, en vertu du principe de la déduction, on peut donc poser $\overset{2}{A} \rightarrow \overset{2}{B}$ isolément, mais $A \rightarrow \overset{2}{A}$ en vertu du principe 6, alors le principe du syllogisme permet d'écrire $A \rightarrow \overset{2}{B}$, de sorte que

$$(d) \qquad \left(\overset{1}{B} \rightarrow \overset{1}{A}\right) \rightarrow \left(A \rightarrow \overset{2}{B}\right);$$

mais on ne peut plus, comme dans la logique formelle, parvenir à la relation symbolique de la démonstration par l'absurde

$$(e) \qquad \left(\overset{1}{B} \rightarrow \overset{1}{A}\right) \rightarrow (A \rightarrow B),$$

le principe $\overset{2}{B} \rightarrow B$ nous faisant défaut.

La démonstration par l'absurde n'est donc plus valable et les deux implications $A \rightarrow B$ et $\overset{2}{B} \rightarrow \overset{1}{A}$ ne sont pas équivalentes.

Mais supposons que la proposition B exprime elle-même l'absurdité de quelque autre proposition $B \equiv \overset{1}{C}$. Alors $\overset{2}{B} \equiv \overset{3}{C}$, mais $\overset{3}{C} \to \overset{1}{C}$, de sorte que, dans ce cas, $\overset{2}{B} \to B$ et l'on passera sans peine par les principes 1 et 2 de la relation (d) à la relation (e). En mettant en évidence la proposition C dans la relation (e), celle-ci s'écrit :

$$\left(\overset{2}{C} \to \overset{1}{A}\right) \to \left(\overset{}{A} \to \overset{1}{C}\right)$$

et c'est bien une forme spéciale de la démonstration par l'absurde.

Les deux implications $A \to \overset{1}{C}$ et $\overset{2}{C} \to \overset{1}{A}$ sont d'ailleurs équivalentes.

La démonstration par l'absurde est donc valable si la proposition que l'on veut établir exprime l'absurdité de quelque autre proposition. En d'autres termes :

On peut prouver l'absurdité d'une proposition par une démonstration par l'absurde.

Il est curieux de constater que, si l'on a le droit de faire de l'absurdité d'une proposition une nouvelle proposition, l'inverse n'est pas vrai, en ce sens qu'on ne peut faire de toute proposition l'absurdité de quelque autre.

Dans la logique formelle, au contraire, une proposition quelconque équivalait toujours à la négation de sa négative.

Les propositions qui représentent l'absurdité de quelque autre proposition ont une signification purement logique que n'a pas une proposition quelconque.

Ainsi, dans notre exemple du début, l'affirmation — le nombre x est algébrique, — proposition traduisible par une équation effectivement construite a une autre signification que l'affirmation — le nombre x est transcendant, — proposition purement logique qui signifie qu'une telle équation serait absurde.

Les démonstrations par l'absurde ne nous apprenaient, en somme, ni le pourquoi ni le comment; elles répondaient comme le sphinx : oui, ou bien : non; qu'elles ne soient plus valables en logique empiriste, où l'on exige le comment, il ne faut pas s'en étonner.

Voici un exemple donné par M. Brouwer; nous l'avons placé à la fin de cette étude à cause de son caractère technique.

Disons qu'un nombre g est rationnel si l'on peut déterminer deux entiers p et q tels qu'on ait $g = \frac{p}{q}$ et irrationnel si l'hypothèse de sa rationalité peut être réduite à l'absurde.

Définissons ensuite un nombre r de la manière suivante : soit d_ν le $\nu^{\text{ième}}$ chiffre, après la virgule, du développement décimal du nombre π. Si, en parcourant ce développement, les 10 décimales consécutives $d_m, d_m + 1, d_m + 9$ forment la suite 0, 1, 2, ..., 9 pour la $\nu^{\text{ième}}$ fois, introduisons un nombre k_n au moyen de l'égalité $k_n = m$. Posons encore $c_\nu = \left(-\frac{1}{2}\right)^{k_1}$ pour $\nu \geq k_1$, sans quoi $c_\nu = \left(-\frac{1}{2}\right)^{\nu}$; soit r la limite des nombres $c_1, c_2, c_3, \ldots$.

Et M. Brouwer d'ajouter : le nombre r n'est ni rationnel ni irrationnel. Cette affirmation arrête parfois ceux qui sont les mieux disposés à admettre son point de vue. Car il est fort possible que l'on parvienne un jour à prouver l'existence et à déterminer le nombre k_1 et alors r sera rationnel. Et M. Brouwer dirait, si je ne fais erreur, r est devenu rationnel; la notion d'un devenir mathématique apparaît souvent dans ses écrits. Certes, on peut dire cela; mais remarquons que la question de savoir si un nombre était rationnel avant qu'on ait démontré qu'il l'est effectivement est aussi de celles qu'il faut renvoyer aux philosophes.

Pour ma part, je dirais simplement : on ne sait si r est rationnel et l'on ne voit pas qu'il doive l'être, quoiqu'il ne puisse être irrationnel. Pour moi, M. Brouwer fournit l'exemple de l'éventualité d'un tertium, mais pas le tertium lui-même, et il serait vain de chercher à faire plus.

Cet exemple sert à plusieurs usages. En effet, comme le remarque l'éminent mathématicien hollandais :

Le nombre r ne peut être proclamé rationnel quoique son irrationalité soit absurde; on ne peut affirmer ni l'absurdité ni, d'autre part, l'absurdité de l'absurdité de l'existence du nombre k_1; on ne peut dire ni que l'ensemble des nombres k est fini, ni d'autre part, qu'il est infini; on ne peut dire, enfin, ni que le nombre r soit supérieur à zéro, ni qu'il soit inférieur à zéro, ni qu'il soit égal à zéro.

Et j'ajouterai : c'est parce qu'aucune de ces propositions particulières ne peut être affirmée aujourd'hui, qu'il est douteux que l'une au moins doive être vraie.

Conclusion. — Le formalisme est à l'abri de la contradiction, me semble-t-il, et il serait fort regrettable qu'il fût abandonné. Mais peut-on faire du non contradictoire le vrai ? Comme le dit M. Brouwer, un homme peut être coupable sans qu'un tribunal l'ait condamné. A-t-on le droit d'accorder même valeur à l'existence idéale qu'à l'existence empiriste ? (principe qu'il faudrait rapprocher de celui de l'induction complète). D'autre part, l'extrême prudence des empiristes me fait songer à la situation où nous nous serions trouvés si nous avions renoncé à la notion du temps absolu, avant que l'expérience se soit prononcée, et même plus, en étant convaincus qu'aucune expérience ne pourrait jamais nous obliger à reviser cette notion.

Rolin Wavre.

II. — Sur le principe du tiers exclu et sur les théorèmes non susceptibles de démonstration (1).

Un récent article de M. R. Wavre a attiré l'attention des lecteurs de cette *Revue* sur les travaux de M. Brouwer relatifs au principe du tiers exclu.

(1) *Revue de Métaphysique et de Morale,* avril 1926.

L'objet du présent article est, d'une part, de montrer qu'il y a, à propos de l'exemple de M. Brouwer cité à la fin de l'article de M. Wavre, un malentendu provenant des définitions employées; d'autre part, d'insister sur ce qui me paraît le point essentiel du débat.

1. Pour faire comprendre le malentendu en question, je m'expliquerai à propos d'un exemple particulièrement simple. Je suppose qu'on ait défini les mots *pair* et *impair* de la manière suivante : Un nombre entier n (positif ou nul) est pair si *l'on peut trouver* un autre entier dont il soit le double; il est impair si $n - 1$ est pair.

Ceci fait, je suppose que quelqu'un tienne le raisonnement suivant : « Le numéro gagnant au premier tirage du Crédit Foncier qui aura lieu en 1930 est sans doute un entier bien déterminé; mais on ne le connaît pas; *on ne peut donc pas trouver* un entier p tel qu'il soit égal à $2p$ ou $2p + 1$; il n'est donc ni pair ni impair. Sans doute deviendra-t-il un jour pair ou impair; mais on peut donner de même des exemples d'entiers qui ne seront jamais ni pairs ni impairs; il suffit de les prendre dans un passé dont l'histoire ne pourra sans doute jamais être reconstituée avec une précision suffisante. Par exemple, le nombre des personnes arrivées en Égypte sur le bateau de Cléopâtre après la bataille d'Actium n'est et ne sera jamais ni pair ni impair. »

Qui ne voit pas qu'il n'y a pas là autre chose qu'un malentendu provenant d'une définition mal exprimée ? Disons franchement qu'à notre avis il n'y a pas autre chose qu'un jeu de mots.

Une définition des mots *pair* et *impair* qui ne donne pas lieu aux mêmes difficultés peut être exprimée de la manière suivante : un nombre entier n est pair s'*il existe* un entier dont il soit le double; il est impair si $n - 1$ est pair. D'ailleurs, la manière même dont on forme la suite des nombres entiers montre clairement que tout nombre entier est pair ou impair, aucun autre cas n'étant concevable; cela résulte clairement aussi de la théorie de la division. Si quelqu'un conteste ce résultat, j'estimerai que son cerveau n'est pas fait comme le mien, et que toute discussion est inutile.

Si maintenant, l'on se reporte à l'exemple cité par M. Wavre, on verra que la difficulté provient de ce qu'il appelle rationnel un nombre égal au rapport de deux entiers p et q *que l'on puisse déterminer*. Après cette définition, il cite l'exemple d'un nombre évidemment rationnel au sens ordinaire de ce mot, mais égal au rapport de deux entiers que, dans l'état actuel de la science, on ne peut pas déterminer, et que peut-être on ne pourra jamais déterminer; ce n'est que par suite d'une mauvaise définition du mot rationnel que l'on peut dire qu'un tel nombre n'est pas rationnel.

Sous ce paradoxe se dissimule, d'ailleurs, un problème intéressant, clairement posé par M. Wavre dans le reste de son article. C'est de savoir s'il est possible, soit de la manière indiquée par M. Brouwer, soit d'une autre manière, de citer un nombre défini sans ambiguïté par une définition mathématique parfaitement claire, et que, cependant, l'on ne puisse pas calculer effectivement. En d'autres termes, est-il possible de citer un problème mathé-

matique parfaitement bien posé, dont la réponse soit *oui* ou *non* sans ambiguïté possible et sans tierce hypothèse logiquement possible, mais tel que les règles de la logique soient impuissantes à déterminer jamais laquelle des deux éventualités est réalisée ? Voilà le problème essentiel, qu'un esprit habitué à la critique philosophique ne peut pas ne pas se poser. Mais, à notre avis, on ne peut que masquer la nature véritable de ce problème en fabriquant des paradoxes comme celui de M. Brouwer.

2. D'où vient qu'un théorème pourrait être vrai et non démontrable ? La raison en est bien simple. Un théorème est en général un énoncé qui comprend une infinité de cas particuliers; chacun de ces cas comporte une vérification. Mais il peut arriver, par des procédés dont nous n'analyserons pas ici le mécanisme logique, que toutes ces vérifications puissent être réduites à une même forme. Un certain raisonnement, dans lequel figurera par exemple un entier arbitraire, donnera successivement toutes les vérifications en question si l'on donne à cet entier toutes les valeurs possibles.

Ce qu'il faut admirer, c'est la puissance de l'analyse mathématique qui arrive ainsi, dans tant de cas, à réduire une infinité de vérifications à un raisonnement unique. Qui peut s'étonner qu'elle n'y soit pas parvenue dans tous les cas ? Non seulement cela n'a rien d'étonnant, mais il est *a priori* assez probable qu'il existe certains énoncés, qui résument ainsi en une formule unique une infinité de cas particuliers, et pour lesquels il est impossible de jamais réduire toutes les vérifications nécessaires à un nombre fini d'opérations. On aura alors un théorème *vrai, mais non démontrable.*

Il faut bien remarquer que cela ne veut pas dire qu'on sache jamais démontrer qu'il n'est pas démontrable. Mais à ce sujet il faut faire une distinction importante, qui ne semble pas avoir été signalée. Pour en montrer l'importance, disons tout de suite que : *il est possible que le théorème de Fermat soit indémontrable, mais on ne démontrera jamais qu'il est indémontrable. Au contraire, il n'est pas absurde d'imaginer qu'on démontre qu'on ne saura jamais si la constante d'Euler est algébrique ou transcendante.*

3. Examinons d'abord le théorème de Fermat qui exprime l'impossibilité d'une relation de la forme

$$a^p + b^p = c^p, \tag{1}$$

a, b, c étant des entiers positifs, et p un entier plus grand que 2. Ce théorème implique évidemment une infinité de vérifications possibles. Mais, s'il est faux, c'est-à-dire s'il existe au moins un système d'entiers a, b, c et $p-2$ positifs et vérifiant la relation (1), cela peut être établi par une seule vérification heureuse. De plus, en faisant toutes les vérifications l'une après l'autre (en commençant par les systèmes d'entiers pour lesquels $a+b+c+p=6$, puis 7, ...), on est assuré, si une telle relation existe, d'y arriver après un nombre fini de tentatives.

On voit donc qu'il n'y a que trois hypothèses logiquement possibles.

Première hypothèse. — Le théorème de Fermat est faux, et cela est démontrable; on est sûr, dans ce cas, d'obtenir la démonstration, par des vérifications faites méthodiquement, après un nombre fini d'opérations.

Deuxième hypothèse. — Le théorème de Fermat est vrai, et démontrable.

Troisième hypothèse. — Le théorème de Fermat est vrai, mais indémontrable.

Si la deuxième hypothèse est vraie, cela ne veut pas dire que la science humaine soit assez habile pour le démontrer un jour; peut être la longueur des raisonnements nécessaires dépasse-t-elle ce que l'humanité peut espérer atteindre.

Mais, si la troisième hypothèse est vraie, on ignorera toujours qu'elle l'est. Il est absurde, en effet, d'imaginer qu'on puisse démontrer qu'elle est réalisée, puisque par là même on aurait démontré le théorème de Fermat et cette hypothèse serait fausse. Il est donc absurde d'imaginer qu'on puisse démontrer que le théorème de Fermat est indémontrable.

4. Voici maintenant l'exemple d'une proposition comportant deux alternatives possibles, mais telles que l'une comme l'autre ne puisse être vérifiée que par une infinité d'opérations : la constante d'Euler est algébrique, ou bien transcendante.

Étant donnée une équation algébrique, comment fera-t-on pour savoir si elle admet la constante d'Euler comme racine ? On calculera, par exemple, les décimales successives, d'une part, de la constante d'Euler, d'autre part, de la racine de cette équation qui semble pouvoir être égale à cette constante. Si elle n'est pas effectivement égale à cette constante, il viendra un moment où l'on constatera la différence; mais, à aucun moment, on ne sera assuré de l'exactitude rigoureuse; il faudrait une infinité d'opérations pour vérifier que cette constante est racine d'une équation algébrique.

Pour savoir si la constante d'Euler est transcendante ou algébrique, on essayera ainsi toutes les équations algébriques, l'une après l'autre. Mais dans aucune des deux alternatives la vérification ne sera jamais terminée. Si, en effet, cette constante est transcendante, on n'en sera sûr qu'après avoir essayé toutes les équations algébriques; si elle est algébrique, les calculs de vérification conduiront certainement, après un nombre fini de tentatives, à former l'équation qu'elle doit vérifier, mais il faudra un nombre infini d'opérations pour s'assurer que cette équation est vérifiée exactement.

En somme, l'énoncé : « La constante d'Euler ne vérifie aucune équation algébrique à coefficients entiers » est bien, comme l'énoncé du théorème de Fermat, un énoncé qui réunit une infinité de cas particuliers dans lesquels il doit être vérifié, et l'échec d'une seule vérification suffit à établir sa fausseté; mais cet échec ne peut être constaté qu'après une infinité d'opérations. Là est la différence essentielle.

Qu'en résulte-t-il au point de vue qui nous occupe ? Il y a ici quatre cas possibles. Nous avons, en effet, d'abord une double alternative suivant que la constante d'Euler est algébrique ou transcendante; mais dans chacun de ces deux cas, le théorème peut être démontrable ou indémontrable. Cela fait, en tout, quatre hypothèses, à première vue possibles toutes les quatre.

Il est alors possible que l'on démontre un jour que le problème de savoir si la constante d'Euler est algébrique, est un problème insoluble. Comme cette circonstance est logiquement compatible avec les deux alternatives que pose le problème, on peut montrer qu'elle se produit sans avoir par là même résolu le problème, comme c'est le cas pour le théorème de Fermat.

Au point de vue des progrès possibles de la science dans cet ordre d'idées, cela laisse encore place à un assez grand nombre d'hypothèses possibles. On distingue, en effet, les quatre cas suivants :

a. La constante d'Euler est algébrique, et on peut le démontrer.

b. La constante d'Euler est transcendante, et on peut le démontrer.

c. Le problème est insoluble, c'est-à-dire qu'on ne saura jamais si la constante d'Euler est algébrique et transcendante, et on peut démontrer qu'il est insoluble.

d. Le problème est insoluble, mais on ne peut pas démontrer qu'il est insoluble.

Naturellement, si cette dernière hypothèse est vraie, elle ne permet pas d'espérer qu'on sache jamais qu'elle est vraie. Cela n'exclut pas pour cela tout progrès ultérieur de la science. On peut, par exemple, espérer trouver un moyen de vérifier sûrement, par un nombre limité d'opérations, si la constante d'Euler vérifie ou non une équation algébrique donnée. Alors on se trouvera réduit à une alternative analogue à celle du théorème de Fermat; l'alternative *c* sera exclue et l'on restera dans l'hésitation entre les alternatives *a*, *b* et *d*.

5. Nous venons de voir qu'il peut exister deux types de théorèmes indémontrables. Nous ne savons pas s'ils existent effectivement, mais il nous paraît logiquement possible qu'ils existent. Il serait intéressant de le démontrer. Il serait plus intéressant encore d'en donner un exemple précis, qui serait nécessairement du second type. Au sujet du premier type, on ne peut en donner un exemple précis, mais on peut espérer trouver un ensemble d'énoncés de ce type tels que l'on sache que l'un d'eux est indémontrable.

Mais il nous semble que la logique a rempli son rôle, en marquant des cases que certains théorèmes viendront peut-être remplir un jour. La parole est aux mathématiques, et la tâche qui leur incombe n'a pas l'air facile.

PAUL LÉVY

III. — Sur le principe du tiers exclu [1].

C'est un très beau médaillon classique que M. Paul Lévy a dessiné sur l'écrin romantique que j'avais essayé de tendre.

Mais nul ne s'étonnera que je relève que pour M. Lévy le point essentiel du débat n'est pas exactement le même que pour moi. Je crois pouvoir le dire sans que M. Lévy trouve trop compromettante l'expression de mon amitié et de mon grand respect.

Le paradoxe du Crédit Foncier ou celui du bateau de Cléopâtre ne me retiendront pas. Le problème relève de l'histoire et non des mathématiques, et l'on ne verrait pas bien ce que j'aurais à faire en cette galère.

Nous sommes d'ailleurs tous d'accord qu'un nombre entier est pair ou impair; il suffit, en effet, de connaître son dernier chiffre pour trancher l'alternative d'un clin d'œil en faveur de l'une ou de l'autre de ses parties.

La difficulté surgit lorsque M. Paul Lévy nous propose de proclamer un nombre pair *s'il existe* un entier dont il soit le double. Plus exactement, elle ne surgit pas encore à propos de la parité d'un entier, mais elle se présente à propos de la rationalité comme aussi à propos de l'algébricité d'un nombre réel.

En effet, pour nous, on ne prouvera la rationalité d'un nombre qu'en découvrant deux entiers dont il soit le quotient et l'on ne démontrera son irrationalité qu'en démontrant que cette découverte impliquerait contradiction. Ces deux circonstances sont seules parfaitement claires. En dehors de ces deux cas, qui ne sont pas les seuls possibles (je crois que M. Lévy le reconnaît) pour moi, on ne sait rien de certain et il ne faut rien préjuger; tout au plus pourrait-on démontrer qu'il serait contradictoire que le nombre fût irrationnel, mais on ne pourrait en conclure qu'il fût rationnel.

Dans la région intermédiaire il faut se défier des mots qui n'étreignent souvent que des ombres, de l'expression *il existe* dont le sens n'est pas toujours immédiat. Il ne servirait à rien de dire : un nombre est rationnel s'*il existe* une fraction égale à ce nombre, car il faudrait encore trouver cette fraction.

Experimentum solum certificat in talibus.

L'exemple de M. Brouwer est tout à fait remarquable; le développement décimal du nombre en question est théoriquement aussi bien défini que celui de la constante d'Euler.

En voici un autre absolument du même ordre en ce qui nous intéresse et dont le développement décimal se calculera aussi facilement et peut-être plus facilement que celui du nombre π.

Soit n le rang de la première décimale précédée de neuf chiffres consécutifs

---

[1] *Revue de Métaphysique et de Morale*, juillet 1926.

égaux à trois que l'on rencontrerait peut-être en parcourant le développement décimal du nombre π ; posons

$$C_i = \left(\frac{1}{10}\right)^i \quad \text{pour } i < n \qquad \text{et} \qquad C_i = \left(\frac{1}{10}\right)^n \quad \text{pour } i \geq n$$
$$(i = 1, 2, 3, \ldots);$$

et soit C la limite des C_i.

Ce nombre C ne saurait être irrationnel, tout le monde est d'accord; il est cependant douteux qu'il soit rationnel, mais il se peut qu'il le soit. Voilà ce que la prudence nous enjoint de dire.

Si les m premières décimales du nombre π nous sont données, les m premières décimales au moins du nombre C nous sont données également.

Nous sommes donc en possession d'une loi de construction du développement décimal du nombre C. Ce dernier peut être calculé approximativement, avec une erreur aussi petite que l'on voudra. Un nombre réel n'est pas autre chose, pour nous, qu'un procédé permettant de construire effectivement un développement, décimal, binaire ou autre. Nous avons insisté sur cet aspect de la pensée empiriste aux pages 452 et 453 du numéro de juillet-septembre 1924 (1). C'est là la seule manière vraiment satisfaisante de définir un nombre réel avec sécurité et précision. Nous jugeons prudent de rejeter, parce qu'elle manquerait de sens précis, toute définition qui ne permettrait pas de procéder à une telle construction. Nous nous défions des définitions formalistes de nombre, dont on n'est même pas certain de pouvoir jamais obtenir ne fût-ce que le premier chiffre. Voilà pourquoi, dans la théorie empiriste, un entier convenablement défini est pair ou impair, premier ou décomposable. Il est toujours possible, en effet, de trancher ces alternatives par une ou un nombre fini de divisions. Mais nous ne saurions affirmer qu'un nombre réel est nécessairement rationnel ou irrationnel, car nous ne sommes pas certains de pouvoir soit trouver une période dans le développement, soit démontrer qu'il serait absurde qu'il y en eût une. Encore une fois, il ne servirait à rien de dire : il existe une période ou il n'en existe point.

Dans mon précédent article, je crois avoir établi qu'il est vain de vouloir forger un nombre dont on puisse démontrer qu'il n'est ni rationnel ni irrationnel ou bien ni transcendant ni algébrique et fournir de cette manière un véritable tertium. Par conséquent il n'y a aucun espoir d'acculer le formalisme à une contradiction par un exemple de ce genre. On ne court aucun risque de se contredire à raisonner ici suivant les modes classiques. Il serait même fort regrettable que l'on renonçât au formalisme classique, car on lui doit les plus belles démonstrations de l'analyse et de la théorie des ensembles qui ne subsisteraient pas sans lui.

C'est encore au formalisme classique que l'on doit la subtile distinction

---

(1) *Revue de Métaphysique et de Morale.*

de M. Paul Lévy entre les théorèmes indémontrables, et dont je serai toujours le premier à relever toute l'élégance et tout l'intérêt dans notre débat.

Insistons sur cet intérêt.

Je ne sais pas si M. Lévy approuverait les raisonnements suivants auxquels j'applique d'ailleurs la logique classique.

Pour démontrer qu'une proposition est indémontrable il *faut* :

1° Prouver qu'il ne pourrait être contradictoire d'admettre que la proposition fût vraie, de manière qu'il soit possible qu'elle soit vraie;

2° Prouver qu'il ne pourrait être contradictoire d'admettre que la proposition fût fausse, car, dans le cas contraire, s'il pouvait être contradictoire d'admettre que la proposition fût fausse, on pourrait espérer prouver qu'elle est vraie, ce qui doit précisément être exclu.

Mais, alors, la proposition pourrait être admise vraie ou bien fausse, indifféremment, sans qu'on risque jamais d'être contredit; par conséquent on ne pourrait ni prouver qu'elle est vraie, ni prouver qu'elle est fausse, autrement on pourrait être contredit.

Notre condition est donc aussi *suffisante*.

Il faut avouer que de telles propositions feraient fort bien figure de postulats; en bonne axiomatique on agirait envers elles comme envers le postulatum d'Euclide. Ce dernier, d'ailleurs, en est une et *il n'y a que les postulats qui répondent à la question* de M. Lévy : *Existe-t-il des propositions dont on puisse démontrer qu'elles sont indémontrables ?*

A moins d'être très réaliste on ne proclamerait de telles propositions ni vraies ni fausses. Elles marqueraient des points de bifurcation de la science.

La distinction si subtile de M. Paul Lévy nous montre, alors, qu'on ne saurait faire un postulat du théorème de Fermat, mais qu'on aura peut-être un jour le droit de postuler la transcendance de la constante d'Euler et de postuler, par conséquent, aussi, son algébricité.

Il faut en conclure qu'il y a probablement plus de postulats qu'on ne le croit d'ordinaire dans les mathématiques classiques.

Il faudra bien que les questions suivantes soient aussi prises un jour en considération.

Une alternative s'impose-t-elle s'il n'existe aucun moyen de la trancher ?

Pourquoi s'imposerait-elle alors que ni l'une ni l'autre de ses parties ne s'imposeront jamais ? Et l'alternative entre le vrai et le faux n'échappera pas à cet examen.

Une proposition parfaitement claire doit-elle être proclamée vraie ou fausse *a priori* alors qu'il n'existerait aucun moyen de démontrer soit sa vérité, soit sa fausseté ?

Voilà pour nous le cœur du débat.

Certes, M. Lévy est libre de pousser le réalisme jusqu'au point d'admettre qu'une proposition puisse être vraie sans être démontrable et *a fortiori* sans être vérifiable. Poincaré disait, en effet :

« Les hommes ne s'entendent pas parce qu'ils ne parlent pas tous la même langue et qu'il y a des langues qui ne s'apprennent pas.

» Et pourtant, en mathématiques, ils ont coutume de s'entendre; mais c'est justement grâce à ce que j'ai appelé les vérifications; elles jugent en dernier ressort et devant elles tout le monde s'incline. Mais là où ces vérifications font défaut, les mathématiciens ne sont pas plus avancés que de simples philosophes. Quand il s'agit de savoir si un théorème peut avoir un sens sans être vérifiable, qui pourra juger puisque par définition on s'interdit de vérifier ? »

Cette conclusion de Poincaré nous permet de rester sur nos positions respectives; aussi ne vais-je plus chercher à convaincre mais simplement à agréer.

Voici quelques remarques à bâtons rompus. L'illustre mathématicien disait encore que, pour bien poser un problème, il fallait l'avoir déjà résolu. J'aimerais à pouvoir traduire : pour bien poser une alternative, il faut être certain de pouvoir la trancher.

Tant que Zénon d'Élée me menace de sa flèche, je ne puis considérer le problème des rapports de la logique et du temps comme résolu, et je salue sympathiquement M. Brouwer qui cherche à le bien poser.

Aristote, d'ailleurs, se refusait à appliquer le principe du tiers exclu aux futurs contingents, et ne craignait pas de dire :

« Il n'est ni vrai ni faux qu'il y aura demain combat sur mer. »

Ajoutons que les nuances de pensée sont des plus riches et des plus subtiles qui vont de l'intuitionnisme de M. Brouwer à celui de M. Weyl, de l'empirisme de M. Baire à celui de M. Borel ou à celui de M. Lebesgue (empirisme dont nous avons dégagé les conséquences, ce qui nous a conduit à la mathématique de M. Brouwer), puis, en passant par l'attitude de Poincaré, au réalisme de M. Paul Lévy, à celui de M. Hadamard, au réalisme cantorien devenu le formalisme axiomatique et conventionnaliste de M. Hilbert et jusqu'au réalisme des classes de M. Russell.

Enfin, je formulerai le vœu que M. Brouwer nous expose lui-même son intuitionnisme et son attitude vis-à-vis de la logique aristotélicienne, point de vue des plus révolutionnaires, gros des principaux problèmes de la philosophie, s'il est vrai que ce soit encore entre Héraclite et Parménide que nous nous débattons comme Aristote lui-même.

Rolin Wavre.

IV. — Critique de la logique empirique (1).

1. En lisant le dernier article de M. Wavre, je me félicite de lui avoir donné l'occasion de l'écrire en adressant quelques objections à son premier exposé.

(1) *Revue de Métaphysique et de Morale*, octobre 1926.

Je crois maintenant avoir compris l'attitude qu'il adopte relativement aux règles de la logique empirique; mais je ne saurais le suivre. Je ne puis admettre l'interdiction arbitraire d'énoncer certains résultats qui me paraissent évidents, de poser certains problèmes qui me paraissent se poser, et d'employer dans ce but le langage qui me paraît le plus commode, et qui d'ailleurs est conforme à l'usage. M. Wavre justifie cette interdiction en disant que ce langage ne lui paraît pas parfaitement clair; je cherche en vain ce qu'il y trouve d'obscur.

Tout d'abord, l'idée de la division d'un ensemble d'objets en deux classes complémentaires me paraît bien nette; par définition tout objet de cet ensemble est de l'une ou l'autre classe, et aucun ne saurait être des deux. Tel est le cas de la division des nombres réels en nombres rationnels et irrationnels. On sait par exemple que les nombres rationnels peuvent être rangés en une suite dénombrable

$$a_1, \quad a_2, \quad \ldots, \quad a_n, \quad \ldots,$$

que j'appelle suite S, où chacun d'eux figure une fois et une seule; il y a bien des manières de le faire, et il est inutile d'allonger cet exposé en précisant celle que je choisis. Cette suite une fois définie, on peut dire qu'un nombre x est *rationnel* ou *irrationnel* suivant qu'il existe ou non dans la suite S un nombre égal à x. Tout nombre est donc rationnel ou irrationnel; il y a là un fait objectif, et les difficultés que je puis éprouver à ranger un nombre dans l'une ou l'autre catégorie, et, s'il est rationnel, à dire son rang dans la suite S, ne sauraient l'empêcher d'être exact.

D'où viennent ces difficultés ? De ce que le nombre le plus simple peut se trouver défini par une propriété qui ne permette pas de le reconnaître. Telle intégrale que je ne sais pas calculer peut se trouver égale à l'unité; je n'en saurai rien et le calcul approché que je puis faire ne m'apprendra pas si le nombre qu'elle définit est rationnel ou irrationnel. Il en est de même pour la plupart des nombres définis par une formule mathématique; il a fallu le génie d'Hermite pour découvrir que le nombre π est irrationnel, et l'on n'est pas encore renseigné sur la nature de la constante d'Euler. Tout cela n'a rien d'étonnant, et je ne vois pas pourquoi je m'interdirais de dire à propos de la constante d'Euler ou de tel autre nombre défini par une formule mathématique : *ce nombre est rationnel ou irrationnel, puisqu'il faut bien qu'il soit l'un ou l'autre, mais je ne sais pas lequel des deux.*

Parmi les différentes circonstances qui peuvent se produire, M. Wavre a été particulièrement frappé par la suivante : il existe des formules définissant des nombres sûrement rationnels, mais dont on ne peut pas, dans l'état actuel de la Science, indiquer le rang dans la suite S. (Il est entendu que j'emploie le langage que je trouve commode pour exposer des faits que M. Wavre expose autrement.) Cela est facile à comprendre; il suffit de choisir un nombre dans la suite S, ou dans une suite S' extraite de S, et de faire dépendre ce choix de circonstances sur lesquelles, dans l'état actuel de la Science, on n'est pas exactement renseigné.

Voici maintenant où nous cessons d'être d'accord. Pour M. Wavre, le

nombre ainsi choisi n'est pas rationnel, car on ne peut prouver la rationalité d'un nombre qu'en indiquant son rang dans la suite S. Je conteste formellement ce point; un nombre étant choisi dans cette suite, je n'ai pas besoin de savoir son rang pour affirmer qu'il est rationnel. Je crois même que je ne fais qu'appliquer une des règles les plus classiques de la logique formelle : les nombres de la suite S sont rationnels; le nombre x est dans la suite S; donc le nombre x est rationnel.

M. Wavre ajoute encore : « A quoi cela vous sert-il de dire que ce nombre est rationnel, si vous ne pouvez pas me dire son rang dans la suite S ? » Je réponds : « Il ne s'agit pas de savoir si cet énoncé est utile, mais s'il est exact, ce dont je ne puis douter. C'est un problème de savoir si le nombre x est rationnel, et je puis le résoudre par l'affirmative; c'en est un autre de savoir son rang dans la suite S, et je ne l'ai pas résolu. Je comprends que la solution de ce second problème vous intéresserait davantage; je vous sais gré de m'avoir signalé son intérêt; mais ce que je ne puis admettre, c'est que sous prétexte que je ne l'aie pas résolu vous m'interdisiez de dire ce que je sais sur le premier. L'attitude que j'adopte, de dire ce que je sais dans un langage que tout le monde peut comprendre, et d'avouer ce que j'ignore, me paraît plus scientifique que la vôtre. »

Généralisant le problème, M. Wavre demande ensuite : « Une alternative s'impose-t-elle, s'il n'existe aucun moyen de la trancher ? » Si je ne me trompe, cela revient à dire : « Peut-on poser un problème sans être sûr de pouvoir le résoudre ? » Je ne crois pas me tromper, puisque M. Wavre rappelle que, d'après Poincaré, « pour bien poser un problème, il faut déjà l'avoir résolu ».

Là encore, il me semble que M. Wavre ne tire pas, de remarques fort justes, la conclusion qui s'impose. Il est certain que bien souvent des problèmes ont été mal posés et qu'il a fallu par la suite en modifier les énoncés. Je n'en citerai qu'un exemple : Michelson a cherché à mettre en évidence le mouvement relatif de l'éther et du système solaire; il a fallu par la suite reconnaître que le problème n'avait pas de sens; mais c'est tout de même de ce problème qu'est née la théorie de la relativité.

Quelle attitude convient-il donc d'adopter ? Pour ma part, je ne reconnais pas au philosophe le droit d'arrêter les progrès de la Science en interdisant au savant d'étudier un problème s'il n'est pas dès le début sûr de le bien poser, et, d'ailleurs, la Philosophie a profité autant que la Science des recherches de Michelson sur un problème mal posé. Le philosophe peut seulement examiner les problèmes que les savants énoncent, se demander s'ils sont bien posés, et s'il n'y a pas dans les raisonnements une fissure rendant possibles d'autres solutions que celles que les savants ont prévues. Il me semble, en prenant la question à ce point de vue, que, même avant les travaux d'Einstein, un philosophe aurait pu observer que l'on ne savait pas ce que c'était que l'éther, et que les raisonnements reposant sur l'assimilation de l'éther et d'un fluide ordinaire pouvaient conduire à des conclusions inexactes. Mais il s'agit de trouver la fissure, et lorsque je suis en présence d'une alternative aussi nette que celle de savoir si un nombre est rationnel ou irrationnel, il

ne suffit pas de me citer l'exemple de problèmes bien plus délicats où les plus grands savants se sont trompés pour que je m'interdise de poser la question ou que j'admette qu'on peut échapper à une alternative à laquelle par définition même on ne peut pas échapper.

On voit la différence irréductible des points de vue. Le fait que M. Wavre parle un langage différent du mien, et n'appelle rationnels que les nombres dont il sait le rang dans la suite S, n'a qu'une importance secondaire. Il distingue ainsi les nombres rationnels, les nombres irrationnels, et ceux qui ne sont ni rationnels ni irrationnels. Je ferai remarquer que, telle intégrale que je ne sais pas calculer pouvant se trouver égale à l'unité, le nombre 1 peut se trouver rangé dans la première ou la troisième catégorie suivant la phrase que j'emploie pour le définir. A une classification des nombres, M. Wavre substitue une classification des phrases dont nous nous servons pour définir les nombres. Il est d'ailleurs parfaitement conséquent avec lui-même en agissant ainsi; mais pour ma part je désire, dans la mesure où cela est possible, mettre en évidence le caractère objectif des choses, et distinguer leurs propriétés objectives des circonstances qui dépendent de l'état de nos connaissances. Le langage classique me convient dans ce but, et je n'aime pas un langage qui, projetant en quelque sorte notre ignorance sur les faits eux-mêmes, arrive à masquer ce que nous savons.

2. Je reviens maintenant en arrière pour examiner une question soulevée par mon précédent article et la réponse de M. Wavre. Pour montrer que la notion d'un nombre qui, au point de vue qu'il adopte, n'est ni rationnel ni irrationnel, n'a rien de surprenant, j'avais cherché à former d'autres exemples plus simples encore que les siens, et à montrer que dans le même ordre d'idées on pourrait même parler de nombres entiers qui ne seraient ni pairs ni impairs.

L'idée essentielle est toujours la même; je fais dépendre le choix de ce nombre d'un problème qui n'est pas résolu, et, si possible, d'un problème insoluble. Faute de connaître actuellement un problème mathématique dont je puisse affirmer qu'il est insoluble, j'avais pris un exemple se rattachant au type suivant : j'avais hier une collection d'objets sous la main; je n'ai pas songé à les compter et personne ne l'a fait, et il n'existe plus aucun moyen de connaître leur nombre; ce nombre est donc un entier dont on ne sait pas la place dans la suite des nombres entiers, dont on ne peut dire s'il est pair ou impair, et qui même, au sens de M. Wavre, ne serait ni l'un ni l'autre, et même pas rationnel Pour rentrer dans le domaine des mathématiques, considérons le problème actuellement irrésolu de la nature rationnelle ou irrationnelle de la constante d'Euler C, et l'exemple indiqué par M. Borel du nombre $f(C)$, $f(x)$ désignant une fonction ayant une valeur si x est rationnel et une autre si x est irrationnel (par exemple 0 et 1).

On observera, et c'est pour faire cette remarque que M. Borel avait indiqué cet exemple, qu'il s'agit là d'un nombre dont je ne puis donner une valeur approchée, comme dans le cas d'une intégrale définie dont je ne connais pas

la valeur exacte mais que je puis toujours calculer avec autant de précision qu'il le faut; je ne sais pas si ce nombre est égal à 0 ou 1. Les nombres de MM. Brouwer et Wavre sont au contraire des nombres dont on peut calculer autant de décimales que l'on veut.

Pour M. Wavre, un nombre tel que $f(C)$ ne saurait être considéré comme *défini*; pour moi, je n'aime pas ce langage, qui ne me paraît pas conforme au sens usuel du mot *défini;* mais je reconnais tout l'intérêt de la distinction signalée par M. Borel, et je dirai avec cet auteur que $f(C)$ n'est pas un *nombre calculable.*

Cette différence de langage révèle une différence réelle entre nos conceptions. Je n'admets pas qu'on me dise : « Je considère le nombre $f(C)$ comme n'existant pas, parce que je ne peux pas calculer sa valeur; je vous interdis de raisonner à son sujet, et, même si vous me démontrez que ce nombre a une certaine propriété, je douterai de l'existence de nombres ayant cette propriété. » Quelque grand que soit l'intérêt de cette notion de nombre calculable, je n'admets pas son introduction dans une question où elle n'a rien à voir, et contester mon raisonnement parce que $f(C)$ n'est pas un nombre calculable me paraît absolument arbitraire. Au point de vue du fait que l'on ne saurait échapper à une alternative nettement posée, l'exemple de $f(C)$, nombre entier dont je ne sais pas le rang dans la suite des nombres entiers, et le nombre de M. Wavre, nombre rationnel dont je ne sais pas le rang dans la suite des nombres rationnels, jouent exactement le même rôle, et j'affirme de la même manière que $f(C)$ est entier et que le nombre de M. Wavre est rationnel. Ceci dit, bien entendu, ne m'empêche pas de reconnaître tout l'intérêt de la remarque que le nombre de M. Wavre, nombre rationnel dont je ne sais pas le rang dans la suite S, est un nombre calculable, tandis qu'un nombre entier dont je ne sais pas le rang dans la suite des nombres entiers ne saurait être un nombre calculable.

3. Je viens d'examiner deux règles de la logique empirique, qui ne me paraissent être autre chose que des interdictions arbitraires; l'une est l'interdiction de poser un problème que l'on n'est pas sûr de pouvoir résoudre; l'autre est l'interdiction de parler d'un nombre si l'on ne sait pas le calculer. Je voudrais encore indiquer que c'est au même point de vue qu'à mon avis il faut examiner les discussions soulevées par l'axiome de M. Zermelo.

Ces discussions ont mis en évidence qu'il existe des collections d'objets dans lesquelles il est impossible d'en choisir un particulier pour le distinguer des autres sans ambiguïté possible. L'exemple le plus simple est fourni par l'étude des nombres compris entre 0 et 1; ceux qu'on peut définir sans ambiguïté par une phrase d'un nombre fini de mots constituent un ensemble dénombrable E (les remarques faites par Poincaré à propos de l'antinomie de Richard n'empêchent pas cet énoncé d'être finalement exact); l'ensemble complémentaire E' est par définition un ensemble dans lequel on ne peut pas désigner explicitement un nombre particulier. On ne peut pourtant, sans aboutir à d'étranges contradictions, dire que l'ensemble E' n'existe pas

et que l'ensemble dénombrable E comprend tous les nombres de l'intervalle (0,1).

C'est un grand service que M. Borel et l'école empiriste ont rendu à la Science de signaler tout l'intérêt de la distinction entre un objet déterminé et un objet choisi dans une collection suivant une règle que l'on est incapable de préciser. Mais je ne puis les suivre lorsqu'ils voudraient considérer ce dernier objet comme inexistant, et mettre en doute la valeur du raisonnement que je peux faire à son sujet. C'est une troisième interdiction, analogue aux deux précédentes, mais qui me paraît tout aussi arbitraire.

En résumé, l'attitude qui me paraît devoir s'imposer en présence tant des objections de l'école empiriste que des sophismes de Zénon d'Élée, est la suivante : ne pas se soucier du paradoxe de Zénon d'Élée et ne pas se demander si le temps suspend sa marche parce que le philosophe n'a pas le temps d'énumérer tous les événements qui peuvent se produire en une fraction de seconde; ne pas douter d'un raisonnement qui nous paraît parfaitement clair parce qu'il ne vérifie pas certaines conditions arbitrairement énoncées; ne pas douter du résultat acquis parce que nous ne pouvons pas le préciser par un autre résultat que l'empiriste déclare seul intéressant.

Par contre, reconnaître tout l'intérêt des distinctions introduites par l'empiriste, et des précisions qu'il voudrait apporter aux résultats de la science classique.

Paul Lévy.

Note. — Je voudrais encore ajouter, pour répondre à l'article de M. Wavre, deux remarques que je n'ai pas pu faire rentrer dans l'exposé qui précède :

1° Je dis qu'une proposition est indémontrable si l'on ne peut pas montrer que sa fausseté implique contradiction. Ce sont donc deux choses distinctes de dire qu'une proposition est indémontrable et que son contraire est indémontrable; comme je l'ai montré, ces deux circonstances peuvent être réalisées séparément ou simultanément. Cela ne me paraît pas conforme au sens usuel du mot « indémontrable » de ne l'employer que dans ce dernier cas.

2° Je suis absolument d'accord avec M. Wavre sur le fait que, dans ce dernier cas (celui d'une proposition indémontrable dont le contraire est aussi indémontrable), il y a là un *postulat* et, par suite, une *bifurcation*, deux voies étant ouvertes à la Science sans qu'on puisse jamais savoir quelle est la bonne. Par contre, le postulatum d'Euclide, qui n'est pas autre chose qu'une précision apportée à une définition incomplète, m'a toujours paru très mal nommé.

V. — A propos de la récente discussion entre M. R. Wavre et M. P. Lévy (1).

MON CHER DIRECTEUR ET AMI,

Voulez-vous me permettre de vous soumettre quelques réflexions au sujet de la très intéressante discussion entre M. R. Wavre et M. Paul Lévy ?

I. Je voudrais tout d'abord écarter rapidement tout ce qui est relatif à des problèmes plus ou moins bizarres qui n'ont rien de mathématique.

On pourrait définir un nombre irrationnel en disant que chacun de ses chiffres successifs est égal à 0 ou à 1 suivant que la réponse à telle ou telle question est affirmative ou négative.

Il serait d'ailleurs possible de ranger les questions qui peuvent être posées en langue française en les classant d'après le nombre de lettres employées et à nombre égal de lettres par lettre alphabétique comme on le fait dans les dictionnaires. On ne conserverait que celles de ces questions qui comportent une réponse par *oui* ou *non*.

Le nombre ainsi défini donnerait par sa seule connaissance la réponse à toutes les énigmes passées, présentes et futures de la science, de l'histoire et de la curiosité. Si je signale cette fantaisie, c'est simplement pour montrer une fois de plus combien la notion de nombre irrationnel est riche et combien il y a d'illusion dans l'esprit de ceux qui pensent que tous les nombres irrationnels peuvent être *définis*.

Si l'on voulait, d'ailleurs, approfondir la « définition » que je viens de donner, on se heurterait à des difficultés analogues à celles qu'on a souvent signalées sous le nom de paradoxe du transfini, mais il serait relativement aisé d'écarter ces difficultés en utilisant, lorsqu'elles se présentent, les chiffres autres que 0 et 1.

II. Une question plus sérieuse me paraît être celle de savoir si l'on doit considérer un nombre comme virtuellement connu lorsque son calcul théoriquement possible exige cependant un temps et une peine hors de proportion avec les possibilités humaines.

Envisageons, par exemple, le nombre formé par les quatre premières décimales du nombre π, c'est-à-dire 1415. Nous pouvons considérer ensuite le nombre formé par les 1415 premières décimales de π; c'est un nombre de 1415 chiffres dont le calcul n'a pas encore été fait et serait certainement très pénible. Mais nous pouvons continuer la même opération et considérer un nombre de décimales précisément égal à ce nombre de 1415 chiffres. Si nous continuons de la même manière seulement un millier de fois, nous arriverons

(1) *Revue de Métaphysique et de Morale*, juillet 1927.

à définir des nombres dont le calcul pratique n'exigerait pas seulement des myriades d'existences humaines, mais dont l'écriture seule, à supposer qu'ils fussent connus, nécessiterait un poids de papier de beaucoup supérieur au poids du globe terrestre.

Doit-on considérer que le dernier des chiffres du millième nombre ainsi défini est pour nous calculable ?

III. Un certain nombre des exemples donnés par M. Wavre et M. Paul Lévy reviennent à la question de savoir si deux nombres irrationnels plus ou moins difficilement calculables et dont on peut cependant obtenir les chiffres successifs, peuvent être égaux sans qu'il soit possible de démontrer cette égalité.

J'ai déjà fait connaître à plusieurs reprises mon sentiment à ce sujet, notamment dans la Note IV de la nouvelle édition de mes Leçons sur la théorie des fonctions [1]. Je suis convaincu que les progrès des mathématiques conduiront nécessairement à limiter le nombre des opérations au bout desquelles un tel problème peut être résolu. En d'autres termes, si deux nombres définis mathématiquement par des voies différentes se révèlent expérimentalement égaux, en ce sens qu'un nombre suffisant de chiffres décimaux calculés sont les mêmes, il doit être possible de démontrer qu'ils sont effectivement égaux.

Je dois reconnaître que cette affirmation n'est pour le moment qu'un postulat. J'ai donné ailleurs les raisons qui me font croire à sa vérité.

IV. Je n'ai pas encore effleuré le point essentiel du débat, c'est-à-dire le principe même du tiers exclu.

Si je comprends bien la position de la question, il s'agit de savoir s'il est possible que deux affirmations contradictoires puissent être librement énoncées l'une et l'autre, chacune d'elles ne risquant jamais d'entraîner de contradictions, et l'absence de contradiction suffisant à justifier une affirmation. Cela ne me paraît pas douteux lorsqu'il s'agit de problèmes historiques vagues et dépourvus d'intérêt.

Il est certain que si l'on suppose notre calendrier transporté chez les Gaulois, il est entièrement indifférent d'admettre que *Vercingétorix est né un lundi ou un mardi*. Il n'en est peut-être pas de même lorsqu'il s'agit de problèmes mathématiques ; si, par exemple, on se demande si telle décimale du nombre π est un nombre pair ou impair, il me semble difficile de ne pas considérer qu'il y a là une alternative pour laquelle il existe une réponse, même si cette réponse ne peut certainement être jamais obtenue. S'il n'est pas vraisemblable, en effet, que l'on puisse tirer une contradiction de l'affirmation que le dernier chiffre du millième nombre dont nous parlions tout à l'heure est pair ou de l'affirmation que ce dernier chiffre est impair, on peut cependant penser que d'un nombre suffisant d'affirmations de ce genre on pourrait tirer une contradiction.

---

[1] *Voir* ci dessus, Note IV.

Il n'est pas impossible, par exemple, que l'on démontre un jour que dans les chiffres décimaux successifs du nombre π, les dix chiffres sont asymptotiquement aussi fréquents les uns que les autres et que par conséquent il est absurde de supposer que, à partir d'un certain rang, tous les chiffres sont pairs ou tous les chiffres sont impairs.

Si donc on était conduit à admettre qu'il n'y a aucune contradiction à supposer que tel chiffre est pair, il n'y aurait non plus aucune contradiction à faire la même affirmation pour tous les chiffres à partir d'un certain rang, ce qui contredit notre hypothèse.

V. Il reste donc, me semble-t-il, à résoudre la question de savoir s'il peut arriver que l'on démontre effectivement que deux propositions mathématiques contradictoires n'entraînent ni l'une ni l'autre de contradiction.

Pour les raisons que j'ai indiquées au sujet du temps nécessaire pour certaines vérifications, il ne me semble pas impossible d'admettre qu'une telle alternative se pose même pour une question comme le théorème de Fermat malgré les remarques fort correctes de M. Paul Lévy. On serait tour à tour amené à considérer deux postulats contradictoires et à développer les conséquences de chacun de ces postulats. Je me demande, toutefois, si cette vue n'est pas purement théorique. Les difficultés qui s'opposent à la démonstration de propositions comme le théorème de Fermat ou comme la transcendance de la constante d'Euler sont, en effet, le caractère en quelque sorte isolé de ces propositions. Il ne semble pas qu'il soit actuellement possible de tirer une longue série de déductions soit du fait que ces propositions sont vraies, soit du fait qu'elles sont fausses. Du moins, si de telles déductions sont toujours formellement possibles, elles ne sont pas intéressantes en ce sens qu'elles n'ajoutent rien de réellement nouveau à la proposition dont on est parti; elles ne la relient pas à d'autres branches de l'arbre mathématique.

On voit que, dans cette question comme dans beaucoup d'autres, on est conduit à adopter un point de vue subjectif et à ne pas considérer les mathématiques comme une science purement objective, suivant la tendance naturelle des logiciens. Les mathématiques ne sont pas seulement une collection de déductions logiques, pas plus que l'arithmétique n'est une collection de calculs numériques exacts. Ce qui importe, c'est l'ordre et le choix; et, s'il est difficile de définir ce que doivent être cet ordre et ce choix, les mathématiciens s'entendent en général fort bien à ce sujet. Parfois, certaines recherches suscitent un intérêt momentané, mais elles tombent dans l'oubli quelques années plus tard et l'on n'en parle plus; d'autres, au contraire, subsistent et deviennent classiques.

Lorsque l'on a des mathématiques cette notion vivante, on est obligé de reconnaître que certains des exemples cités par M. Wavre et par M. Lévy sont autant en dehors des mathématiques que le nombre fantaisiste dont je parlais au début de cette lettre. En forçant un peu ma pensée, j'arriverais presque à dire que les problèmes insolubles seront abandonnés par les mathé-

maticiens, puisque le fait qu'ils restent insolubles les classe à l'extrémité d'une série de déductions, alors que les questions mathématiques intéressantes se relient les unes aux autres par de nombreuses chaînes qui en font un tissu d'une riche complexité. En fait, on constate que les mathématiciens ne s'intéressent à certains problèmes classiques jusqu'ici sans solution que dans la mesure où l'étude de ces problèmes peut être rattachée à certaines théories générales, théories dont l'intérêt subsiste, même si le problème qui leur a servi de prétexte n'avance pas d'un pas vers sa solution (1). C'est là peut-être la raison profonde pour laquelle les mathématiciens ne se passionnent pas pour ou contre la logique empirique et continuent à creuser patiemment leur sillon.

ÉMILE BOREL.

P. S. — Un mot encore en réponse au reproche que me fait M. Paul Lévy de poser une interdiction arbitraire (p. 278). Je suis beaucoup trop libéral pour avoir jamais songé, dans le domaine de la pensée, à interdire n'importe quelle démarche de l'esprit à aucun savant ou philosophe et à M. Paul Lévy moins qu'à tout autre, car je sais trop quel excellent usage il fait de la liberté à laquelle il a droit, comme chacun de nous. Je n'interdis rien, et je n'interdis pas en particulier à M. Paul Lévy de raisonner sur un nombre α qu'il ne sait pas définir. Mais je constate que M. Paul Lévy ne sait pas mieux que moi-même ce qu'est le nombre α, puisqu'il ne le définit pas. S'il me dit que ce nombre est choisi suivant une règle qu'il conçoit mais qu'il ne précise pas, je veux bien croire qu'il attache un sens à ces mots, mais je ne suis pas plus avancé dans la connaissance du nombre α. Je constate, en outre, et c'est là l'essentiel à mes yeux, que *tout* ce que M. Paul Lévy peut dire du nombre α pourrait être dit exactement sous la même forme d'un autre nombre β choisi, lui aussi, suivant une règle que nous concevrions sans la préciser. En d'autres termes, je constate que M. Paul Lévy, lorsqu'il croit raisonner sur le nombre α, raisonne en réalité sur *tous* les nombres de la collection dans laquelle il a choisi α et étudie par conséquent, non les propriétés de α, mais les propriétés de cette collection, qui a seule une réalité à mes yeux.

Ce que je mets en doute, ce n'est pas la valeur du raisonnement que M. Paul Lévy peut faire sur des nombres tels que α et β, c'est l'utilité et, je dirais, l'existence même de ce raisonnement. Si nous admettons, par exemple, que α et β sont des nombres réels, je ne vois aucun inconvénient à ce que l'on affirme, à propos de α et β, une identité algébrique telle que :

$$(\alpha-\beta)(\alpha+\beta)=\alpha^2-\beta^2. \tag{1}$$

Mais je pense que cette identité n'est utilisable que lorsque α et β y sont connus. L'intérêt qu'il y a à l'écrire sous sa forme générale, c'est qu'on est assuré qu'elle est vraie quels que soient les nombres *déterminés* α et β et qu'elle

(1) C'est le cas pour les célèbres recherches de Kummer sur le théorème de Fermat.

est donc également applicable à des nombres α et β que nous n'avons pas su définir aujourd'hui mais que nous saurons définir demain; M. Paul Lévy, ni tout autre, n'utilisera jamais cette relation pour un calcul sur des nombres α et β qu'il ne connaîtrait pas.

E. B.

VI. — Sur la logique de M. Brouwer [1].

Les polémiques sur le transfini portaient essentiellement sur la notion d'*existence* de certains êtres mathématiques, dont la construction, au sens usuel de ce mot, n'est pas possible, par exemple, l'existence d'une *perordination* (si l'on veut bien nous pardonner ce néologisme, pour traduire Wohlordnung) des ensembles quelconques, notamment du continu.

Or, ce même problème d'existence se pose aussi dans le fini, pour ceux qui acceptent le point de vue de Kronecker, d'après lequel un nombre, un être mathématique, n'*existe* que s'il peut être construit à l'aide des nombres entiers et par un nombre fini d'opérations.

Poussant jusqu'à ses conséquences ultimes cette « arithmétisation » de l'analyse, M. Brouwer a vu clairement qu'elle nous forçait à reviser les principes mêmes de la logique et à rejeter celui du tiers exclu, c'est-à-dire l'affirmation qu'une proposition est toujours vraie ou fausse.

Par exemple, ou bien *tous les nombres ont une certaine propriété a,* ou bien cela est faux et *il existe quelque nombre qui n'a pas cette propriété.* Or, il suffirait de démontrer que la première de ces hypothèses conduit à une contradiction, pour établir l'existence d'un nombre qui n'a pas la propriété *a*. M. Brouwer, n'admettant pas cette démonstration d'existence sans construction, est amené à penser que la fausseté de la première proposition n'entraîne pas forcément la vérité de la seconde; il s'ensuit que celle-ci peut être autre chose que vraie ou fausse.

Il faut donc qu'il y ait un troisième état des propositions et qu'une proposition au moins appartienne à ce *tiers*, c'est-à-dire ne soit ni vraie ni fausse.

Dans notre Note communiquée à l'Académie royale de Belgique, nous démontrons que la notion même d'une telle proposition tierce implique contradiction.

Pour y parvenir, nous sommes partis des postulats suivants, admis d'ailleurs explicitement par M. Brouwer :

1° Principe de double négation : la vérité d'une proposition implique la fausseté de sa fausseté.

2° Principe de transposition : si une proposition en implique une autre, la fausseté de la conséquence implique la fausseté de la prémisse.

---

[1] Résumé d'une Note parue le 8 janvier 1927 dans le *Bulletin de l'Académie royale de Belgique.*

3° Principe de syllogisme : si une proposition en implique une autre, et celle-ci une troisième, la première implique la troisième.

Nous y avons joint quelques postulats purement formels sur le produit et la somme logiques, le produit logique étant une collection de propositions dont chacune est vraie et la somme logique, une collection de propositions dont au moins l'une est vraie :

1° Quand le produit logique est vrai, chacun de ses termes est vrai.
2° Quand un terme d'une somme logique est vrai, la somme logique est vraie.
3° Quand une proposition implique le produit logique de deux autres, elle les implique séparément, et réciproquement.
4° Quand une proposition implique la somme logique de deux autres, elle implique au moins l'une des deux, et réciproquement.
5° Quand deux propositions impliquent séparément une même proposition, leur somme logique implique cette proposition, et réciproquement.

On se rendra compte facilement, tout d'abord, que ces règles sont indispensables au raisonnement, et ensuite, qu'aucune d'elles n'implique le principe du tiers exclu.

Ajoutons encore deux autres postulats :

1° Principe généralisé de contradiction : il est impossible qu'une proposition soit à la fois vraie et fausse, ou fausse et tierce, ou tierce et vraie.
2° Principe du quart exclu, qui n'a jamais été formulé par M. Brouwer, mais qu'il nous paraît impossible de refuser, puisque le tiers est défini d'une façon purement négative, comme étant le « ni vrai ni faux » : une proposition est ou bien vraie, ou bien fausse, ou bien tierce.

Partant de ces postulats, nous établissons qu'ils conduisent à l'une au moins des conclusions suivantes : *Quand une proposition est vraie, elle est tierce* ou *quand une proposition est tierce, elle est fausse.* Or, toutes deux sont absurdes, puisqu'elles violent le principe de contradiction.

Remarquons, pour ceux qu'inquiètent à juste titre, dans une matière aussi délicate, les recours à l'intuition, que notre démonstration est exprimée par les formules d'un calcul symbolique.

Nous venons donc de voir que ces postulats entraînent une contradiction [1]. D'où vient-elle ?

---

[1] Dans une Note récente, M. Paul Lévy a cru justifier la logique brouwerienne, en en fournissant une interprétation empruntée à la logique classique. Nous pensons devoir maintenir nos conclusions; car cette interprétation accepte implicitement deux des postulats, dont le rejet est essentiel au système de M. Brouwer : réciproque du principe de double négation et... principe du tiers exclu !

A notre avis, les postulats avoués par M. Brouwer, principe de transposition et de double négation, en sont responsables, car ils impliquent l'exclusion du tiers; en effet, quand une proposition est vrai, il est clair que sa négociation n'est pas vraie, mais non pas qu'elle soit fausse : elle pourrait être tierce; et de même, quand une proposition en implique une autre, si cette autre est fausse, la première ne peut être vraie : mais elle n'est pas toujours fausse, puisqu'elle aussi peut être tierce.

Logiquement, M. Brouwer ne pouvait donc pas exclure le tiers dans ces deux cas; et en corrigeant les principes comme nous venons de l'indiquer, son système devenait cohérent.

Mais alors, il n'y aurait plus aucun moyen d'établir de rapport entre le vrai et le faux et il nous serait donc impossible de jamais tirer parti de la connaissance qu'une proposition est fausse : il n'existerait plus de relations qu'entre le vrai et le non-vrai, c'est-à-dire le *faux ou tiers*, et ces relations retomberaient sous les règles de la logique classique.

Quoi qu'il en soit, par son échec même, l'œuvre de M. Brouwer a l'immense mérite de démontrer que l'extrême arithmétisation de l'analyse, pour féconde qu'elle apparaisse, n'est qu'une méthode parmi d'autres; qu'une existence non constructive est légitime, même si elle ne présente pas tous les avantages d'une construction; et que le seul critérium universel de validité pour les propositions mathématiques, c'est le respect des règles immuables de la logique.

M. Barzin et A. Errera.

VII. — Logique classique, logique brouwerienne et logique mixte.

1. La très intéressante Note présentée par MM. Barzin et Errera à l'Académie Royale de Belgique m'a donné l'occasion de préciser sur quelques points, dans une Note présentée à cette Académie le 3 mai 1927, les idées que j'ai déjà exprimées sur la logique empirique [1]. Je reproduirai ou résumerai ci-dessous les passages essentiels de cette Note. J'indiquerai d'abord que, si mes résultats sont très différents de ceux de MM. Barzin et Errera, cela tient à la différence de nos points de vue; il n'y a, contrairement à ce que certains passages de ma Note citée peuvent faire croire, aucune contradiction essentielle entre eux et moi.

Leur idée est que, pour montrer le vice inhérent à la logique de M. Brouwer il faut en adopter sans discussion toutes les conventions, et montrer qu'elles conduisent à une contradiction. C'est sans donte ainsi qn'il faut procéder si l'on s'adresse au lecteur déjà troublé par l'argumentation du savant hollandais. Je m'adresse au contraire à « l'homme normal », non infecté du virus brouwerien, croyant encore à son bon sens et à la logique classique, et

(1) *Revue de Métaphysique et de Morale*, 1926.

c'est à ce lecteur que j'essaie, en me servant du langage qui lui est familier, de donner une idée de ce que peut être la logique brouwerienne. M. Brouwer employant manifestement les mots *vrai* et *faux* dans un sens qui n'est pas leur sens ordinaire, j'essaie de définir avec des mots du langage ordinaire les notions qu'il exprime par ces mots, et comme l'on peut hésiter entre deux interprétations légèrement différentes de la pensée de M. Brouwer, j'arrive à définir, à côté de la logique classique, deux logiques nouvelles, que j'appelle *logique brouwerienne* et *logique mixte*.

Il faut bien convenir qu'aucune de ces logiques n'est exactement celle de M. Brouwer, et cela est heureux pour leur développement ultérieur, car MM. Barzin et Errera ont montré que celle-ci aboutit forcément à une contradiction; pour ma part, je ne crains aucun reproche de ce genre. Si j'ai cru bien faire d'introduire dans mon exposé le nom de *logique brouwerienne*, c'est que j'ai été conduit à ces notions en cherchant à tirer une logique viable de ce qui me paraît être l'idée fondamentale de M. Brouwer : *ne tenir pour vrai que ce qui peut être effectivement démontré;* par suite, *n'employer le mot vrai que quand on peut dire « vrai et démontrable »*. Cela est sans doute une précaution élémentaire lorsqu'on dit affirmativement « cela est vrai »; mais un énoncé hypothétique tel que *si cela est vrai* n'implique pas que cette vérité soit démontrable. La logique devant naturellement classer divers cas possibles sans savoir d'avance lequel est réalisé et si l'on peut démontrer qu'il l'est, on est conduit souvent à employer le mot « vrai » dans des cas où il est impossible de dire « vrai et démontrable ». Dans mes logiques nouvelles, après avoir défini tous les cas où l'on peut dire « vrai et démontrable », j'arrive par exclusion à un cas où il faut dire « vrai et non démontrable »; on ne peut éviter de considérer ce cas comme une possibilité logique, bien qu'évidemment l'on ne puisse jamais savoir qu'on est dans ce cas.

Si M. Brouwer accepte ces considérations, il ne peut éviter d'adopter une de mes deux logiques. S'il les refuse, et s'il persiste à appliquer dans tous les cas des règles qui ne sont faites que pour les cas où la vérité implique la possibilité d'une démonstration, il ne peut éviter de les appliquer dans un cas où elles ne s'appliquent pas, et il se condamne d'avance à la contradiction. MM. Barzin et Errera ont eu le mérite de montrer le chemin par lequel il y arrive.

Je n'adresserai à ces auteurs qu'une critique de détail, relative à l'emploi de cet axiome : « *Si une proposition* A *implique la somme logique de deux autres*, B *et* C (c'est-à-dire si elle implique que l'une ou l'autre des propositions B ou C soit exacte), *ou bien elle implique* B, *ou bien elle implique* C. » Je ne comprends pas bien. La proposition « *n est entier* », par exemple, implique « *n est pair ou impair* »; elle n'implique ni « *n est pair*, ni *n est impair* ». L'entier *n*, n'est que quelquefois pair, et l'on ne peut alors considérer que ce soit une conséquence nécessaire de ce que ce nombre soit entier; que signifie alors la phrase « A *implique* B », si elle ne peut vouloir que B résulte nécessairement de A ?

Toutefois, cette objection est moins grave que je ne l'avais pensé d'abord,

car le raisonnement de MM. Barzin et Errera subsiste, à quelques détails près, si l'on remplace leur énoncé par le suivant, évidemment exact : « *Si* A *implique la somme logique de* B *et* C, *et si* C *exclut* A, *alors* A *implique* B ». Leur conclusion finale, relative au vice inhérent à la logique brouwerienne, me semble donc bien exacte.

2. J'arrive à l'exposé de la logique nouvelle.

Considérons un énoncé α, et désignons par α le contraire de β. Des deux énoncés α et β, dans tous les cas possibles, un et un seul est vrai; c'est le *principe classique du tiers exclu.*

Nous dirons que α est *vrai au sens de M. Brouwer* (par abréviation *vrai-Br*; on peut dire aussi *démontrable*), s'il est possible de démontrer la vérité de α. Nous désignerons cette affirmation brouwerienne de α par $+\alpha$, ou encore par A; nous poserons de même $+\beta \equiv$ B. Remarquons que dans tous les cas deux signes $+$ consécutifs équivalent à un seul; s'il est possible de démontrer la vérité de α, il est par là même possible de démontrer celle de A, de sorte que l'on a $+\mathrm{A} \equiv \mathrm{A}$. Nous appellerons proposition brouwerienne ou énoncé brouwerien toute proposition P définie par une formule du type $\mathrm{P} \equiv +p$; on a donc $+\mathrm{P} \equiv \mathrm{P}$; une proposition brouwerienne ne peut pas être vraie sans être par là même démontrable. Ces propositions seront dans la suite désignées par des grandes lettres; les autres par des minuscules grecques.

Des deux énoncés α et β, un et un seul est toujours vrai. Mais des deux énoncés brouweriens correspondants A et B, il peut arriver qu'aucun ne soit vrai; dire qu'on n'a ni A ni B revient à dire que le problème de savoir si α est vrai ou faux n'est pas un problème résoluble. Nous dirons dans ce cas que α est *tiers;* α est donc soit *vrai-Br* (dans le cas A, en appelant A le cas où A est vrai), soit *faux-Br* (dans le cas B), soit tiers, dans un dernier cas, que nous désignerons par α'. Ces trois cas sont exclusifs et seuls possibles. C'est le *principe du quart exclu.*

L'énoncé « α est tiers » n'est pas une proposition brouwerienne; il faut donc distinguer deux cas, suivant qu'on peut ou non démontrer que α est tiers. Si on peut le démontrer, on sera dans le cas $+\alpha' \equiv$ C; si on ne le peut pas, on sera dans le cas α''. Nous retombons ainsi sur un nouveau cas qui n'est pas brouwerien, et cette fois nous ne pouvons pas continuer, car le symbole $+\alpha''$ implique contradiction; si, en effet, on pouvait démontrer qu'on est dans le cas α'', on aurait par là même démontré qu'on est dans le cas α', et par définition du cas α'' on ne peut pas le démontrer. Si donc il est possible *et même nécessaire* de concevoir le cas α'' comme une possibilité logique, on ne saura jamais qu'on est dans ce cas.

On a ainsi quatre cas A, B, C, α'', et il n'est pas possible de les subdiviser à nouveau au point de vue de M. Brouwer. Si par contre on considère la question « α est-il vrai ou faux » comme ayant un sens même dans les cas C et α'', chacun de ces cas se subdivise en deux; C se divise en γ_1 si γ est vrai et γ_2 si α est faux (γ_1 et γ_2 ne sont pas des cas brouweriens, puisque ne sachant pas si α est vrai ou faux on ne peut pas savoir dans quel cas on est); α'' se divise

de même en α_1 et α_2. On arrive ainsi au maximum à six cas A, B, γ_1, γ_2, α_1, α_2, et l'on ne peut pas les subdiviser sans introduire un énoncé indépendant de α; toutes les questions qu'on peut se poser au sujet de α ont dans chaque cas une réponse bien déterminée.

Pour bien montrer l'importance de la distinction entre α et A, demandons-nous dans quels cas A est vrai-Br, ou faux-Br, ou tiers. D'abord A étant une proposition brouwerienne est vrai-Br dans le cas A; il est faux-Br dans le cas B (parce que α est faux-Br) et dans le cas C (parce qu'on peut montrer dans ce cas qu'on est dans le cas α' et par suite pas dans le cas A); il est tiers dans le cas α''; c'est le cas où l'on ne sait rien et où toutes les propositions sont tierces.

3. On remarque que, dans ce qui précède, l'existence des cas C et α'' n'apparaît que comme une possibilité logique. La logique doit considérer comme naturel, en attendant qu'un problème soit effectivement résolu, de prévoir et classer les différents cas qui paraissent possibles; cela ne veut pas dire que tous existent. Il y a quelques raisons de penser que le cas C n'existe pas effectivement et l'existence du cas α'' n'est pas sûre non plus; mais ces raisons sont plutôt psychologiques que logiques, et l'on ne voit guère par quel moyen l'on ferait progresser cette question; il faut donc bien, jusqu'à nouvel ordre, considérer ces cas, et leurs subdivisions γ_1, γ_2, α_1, α_2, comme des possibilités logiques.

Je rappelle d'ailleurs la remarque que j'ai déjà faite que, pour les problèmes actuellement irrésolus, il y a deux cas essentiellement différents. Dans le cas d'un énoncé comme celui ci « *la constante d'Euler est rationnelle* », aucun des six cas A, B, γ_1, γ_2, α_1, α_2 ne peut être exclu, dans l'état actuel de la Science. Au contraire, pour l'énoncé « *le théorème de Fermat est vrai* », les cas A, B, α_1 sont seuls possibles; naturellement, pour l'énoncé contraire, « *le théorème de Fermat est faux* », les cas A, B, α_2 sont seuls possibles.

Je ne crois pas qu'on puisse, sans introduire un principe de raisonnement essentiellement nouveau, faire progresser davantage la question des possibilités effectives des différents cas, soit d'une manière générale, soit pour un énoncé particulier.

4. Revenons maintenant à la logique brouwerienne, et observons qu'il lui est impossible de rester jusqu'au bout fidèle à son idée directrice, qui est de n'énoncer jamais que des propositions brouweriennes, c'est-à-dire telles que leur énoncé implique la possibilité d'une démonstration. Ayant en effet défini les cas A et B, elle ne peut éviter de considérer les cas non brouweriens α' et α''. Mais elle peut éviter de parler de la proposition initiale α, et elle peut sans inconséquence déclarer que cette proposition n'a pas de sens, puisque aucune question relative à la vérité ou à la fausseté (tant au sens ordinaire qu'au sens brouwerien) des cas A, B, C, α'' ne nous conduit à une subdivision de ces cas et ne nous oblige à nous demander si α est vrai.

Mais, si l'on ne veut pas parler de α, comment définira-t-on les cas A, B,

et α' ? La proposition A peut sans doute être considérée comme donnée; mais il faudrait pouvoir ensuite définir B. Or, B est défini comme étant la négation brouwerienne de α, et non de A; nous avons vu plus haut que *la proposition* A *est vraie-Br dans le cas* A, *fausse-Br dans les cas* B *et* C, *et tierce dans le cas* α'' *seulement*. Aucune question relative à la vérité (simple ou brouwerienne) de ces trois cas ou à leur fausseté ne peut nous amener à distinguer B et C; ce sont deux cas où la fausseté de A est démontrable, dans l'un parce qu'on peut démontrer que α n'est pas vrai, dans l'autre parce qu'on peut démontrer que α n'est pas démontrable; peu importe cette distinction, puisqu'on ne s'intéresse qu'à la vérité démontrable de α.

Il n'y a plus alors à distinguer que trois cas, A, l'ensemble des cas B et C que nous désignerons par $\sim$ A (lire *non* A; le symbole $\sim$ désigne ici la négation brouwerienne) [1], et le cas tiers α''. On remarque qu'alors dans cette logique on ne pourra jamais affirmer qu'une proposition déterminée est tierce, le cas α'' n'était jamais démontrable. Cette remarque ne permet pourtant pas de nier qu'on puisse démontrer un jour que la proposition « *la constante d'Euler est irrationnelle* » est tierce, car cette proposition n'a pas droit de cité dans la logique considérée; la seule chose sûre, dans l'état actuel de la Science, est qu'on ne démontrera jamais que la proposition « *on peut démontrer que la constante d'Euler est irrationnelle* » est tierce.

Nous arrivons donc à la conclusion qu'il peut y avoir trois logiques différentes :

1° La logique classique, qui distingue six cas A, B, γ_1, γ_2, α_1, α_2.

2° La logique mixte, qui admet l'énoncé de α pour servir de base à la définition de B et C mais s'interdit de se demander, dans le cas α', si α est vrai ou faux; elle distingue quatre cas A, B, C, α''; sauf lorsqu'il s'agit de α'', elle évite d'employer les mots *vrai* ou *faux* autrement que dans le sens brouwerien.

3° La logique brouwerienne, qui s'interdit absolument de parler de α, considère la notion de A comme une notion première, et ne distingue que trois cas, A, $\sim$ A (remplaçant B et C), et le cas tiers α'', dont l'affirmation brouwerienne n'est jamais possible.

5. *Conclusion*. — Ces différentes logiques, qui ne sont que des branches de la logique classique, sont bien entendu exemptes de contradiction. Mais il est évident qu'on ne peut, ni éviter de considérer le cas α'', défini par l'exclusion des autres cas, ni raisonner sur ce cas comme on le fait sur les propositions brouweriennes, pour lesquelles l'énoncé implique la possibilité d'une démonstration. Les règles de logique établies en vue des raisonnements sur des propositions brouweriennes ne peuvent donc être appliquées dans tous les cas et l'on ne saurait éviter d'aboutir à une contradiction si on les applique

---

[1] Si $\sim$ désigne la négation simple, la négation brouwerienne devra être désignée par $+\sim$.

sans discernement. Les brouweriens ne peuvent donc pas rester jusqu'au bout fidèles à leur idéal.

Pour ma part, je ne saurais le regretter. Un des buts de la logique n'est-il pas, comme je l'ai dit plus haut, de prévoir et de classer les différents cas logiquement possibles, sans savoir lequel est réalisé et même s'il est possible de le savoir ? Cette remarque ne suffit-elle pas à écarter les objections des brouweriens contre le principe du tiers exclu ?

Des travaux des empiristes, je voudrais surtout, comme je l'ai dit dans mon article cité plus haut, retenir cette remarque essentielle que certains énoncés n'ont pas l'aspect absolument précis et tangible auquel l'arithmétique et l'algèbre élémentaires nous ont habitués. Je m'efforce de comprendre la psychologie des empiristes qui les évitent autant que possible, et je voudrais qu'ils comprennent qu'il y a, au contraire, pour l'idéaliste un attrait supplémentaire dans le fait que ces énoncés élargissent le champ des recherches mathématiques; il faut souhaiter que la Science progresse dans différentes directions; chacun, empiriste ou idéaliste, est libre de réserver ses efforts à celle qui l'intéresse. Aucune contradiction logique ne devrait exister entre les deux points de vue; les uns et les autres devraient reconnaître les mêmes faits et se borner à constater la différence de leurs psychologies.

Mais si l'empiriste veut donner une plus grande portée à ses objections, s'il s'agit, non d'une psychologie empirique, mais d'une logique empirique, si l'on conteste la valeur des résultats obtenus par les idéalistes, et si, pour empêcher plus sûrement les progrès de la pensée idéaliste, on crée un langage spécial, changeant le sens de mots usuels tels que *vrai* et *faux* ou d'expressions telles que *l'ensemble des fonctions continues de x définies de* 0 *à* 1, alors je me range à côté de ceux qui ne veulent tenir aucun compte de ces objections.

PAUL LÉVY.

TABLE DES MATIÈRES

NOTES DE LA PREMIÈRE ÉDITION.

NOTES DE LA DEUXIÈME ÉDITION.

NOTE DE LA TROISIÈME ÉDITION.

FIN DE LA TABLE DES MATIÈRES.

Envoi dans toute l'Union postale contre mandat-poste ou valeur sur Paris.
Frais de port en sus (Chèques postaux : Paris 29 323). R. C. Seine 22 520.

Collection de monographies sur la Théorie des fonctions, publiée sous la direction d'ÉMILE BOREL, Membre de l'Institut. Volumes in-8 (25-16) se vendant séparément.

Leçons sur la théorie des fonctions (*Éléments et principes de la théorie des ensembles; applications à la théorie des fonctions*), par EMILE BOREL.

Leçons sur les fonctions entières. 2e édition .... 25 fr.

Leçons sur les séries divergentes, par EMILE BOREL. 2e édition. .. fr.

Leçons sur les séries à termes positifs, professées au Collège de France par EMILE BOREL; recueillies et rédigées par ROBERT D'ADHÉMAR; 1902.. 25 fr.

Leçons sur les fonctions méromorphes, professées au Collège de France par EMILE BOREL, recueillies et rédigées par LUDOVIC ZORETTI; 1903.. 25 fr.

Leçons sur l'intégration et la recherche des fonctions primitives, professées au Collège de France, par HENRI LEBESGUE. 2e édition. (*Sous presse.*)

Leçons sur les fonctions de variables réelles et les développements en séries de polynomes, professées à l'École Normale supérieure, par EMILE BOREL et rédigées par MAURICE FRÉCHET, avec des Notes par PAUL PAINLEVÉ et HENRI LEBESGUE. 2e édition.............................. 25 fr.

Leçons sur les fonctions discontinues, professées au Collège de France, par RENÉ BAIRE, rédigées par *A. Denjoy*; 1905 .... 25 fr.

Le calcul des résidus et ses applications à la théorie des fonctions, par ERNST LINDELÖF; 1905.. 25 fr.

Leçons sur les séries trigonométriques, professées au Collège de France, par HENRI LEBESGUE; 1906.............................. (*Sous presse*).

Leçons sur les fonctions définies par les équations différentielles de premier ordre, professées au Collège de France par PIERRE BOUTROUX, Maître de Conférences à la Faculté des Sciences de Montpellier; avec une *Note* de PAUL PAINLEVÉ; 1908 .... 25 fr.

Principes de la théorie des fonctions entières de genre infini, par OTTO BLUMENTHAL; 1910 .... 25 fr.

Leçons sur la théorie de la croissance, professées à la Fté des Sciences de Paris, par É. BOREL, recueillies et rédigées par *A. Denjoy*; 1910.... 25 fr.

Leçons sur les séries de polynomes à une variable complexe, par PAUL MONTEL; 1910 .... 25 fr.

Leçons sur le prolongement analytique, professées au Collège de France, par LUDOVIC ZORETTI; 1911.............................. 25 fr.

Leçons sur les équations intégrales et les équations intégro-différentielles, professées à la Faculté des Sciences de Rome en 1910, par VITO VOLTERRA, publiées par *M. Tomassetti* et *F.-S. Zarlatti*; 1913............ 25 fr.

Leçons sur les singularités des fonctions analytiques, professées à l'Université de Budapest, par PAUL DIENES; 1913.................... 25 fr.

Leçons sur les fonctions de lignes, professées à la Sorbonne en 1911, par VITO VOLTERRA, recueillies et rédigées par *J. Perès*; 1913... 25 fr.

Les systèmes d'équations linéaires à une infinité d'inconnues, par FRÉDÉRIC RIESZ; 1913.. 25 fr.

TRAITÉ DU CALCUL DES PROBABILITÉS ET DE SES APPLICATIONS

Par Émile BOREL

Professeur de Calcul des Probabilités et de Physique mathématique
à la Faculté des Sciences de Paris,
Membre de l'Institut.

TOME I. — *Les principes de la Théorie des probabilités.*

1. Principes et formules classiques, par Émile BOREL, rédigé par R. LAGRANGE, 160 pages, 29 figures ; 1925 .... 25 fr.
2. Erreurs et moindres carrés, par Robert DELTHEIL.
3. Recherches théoriques modernes, par Maurice FRÉCHET.
4. Les principes de la statistique mathématique, par C.-E. TRAYNARD.

TOME II. — *Les applications de la Théorie des probabilités aux sciences mathématiques et aux sciences physiques.*

1. Applications à l'arithmétique et à la théorie des fonctions, par Émile BOREL, rédigé par Paul DUBREIL ; 1926........... 25 fr.
2. Probabilités géométriques, par Robert DELTHEIL ; 1926... 30 fr.
3. Mécanique statistique classique, par Émile BOREL, rédigé par Francis PERRIN, 148 pages, 13 figures ; 1925 .... 25 fr.
4. Applications à l'astronomie, par C.-V.-L. CHARLIER.
5. Applications aux théories physiques actuelles, par Francis PERRIN.

TOME III. — *Les applications de la Théorie des probabilités aux sciences économiques et aux sciences biologiques.*

1. Assurances sur la vie. Calcul des primes, par Henri GALBRUN, 210 pages ; 1924.................................... 45 fr.
2. Assurances sur la vie. Calcul des réserves, par Henri GALBRUN ; 1927.. 40 fr.
3. Applications à la biologie. Variations discontinues et Mendélisme, par L. BLARINGHEM........................... (*Sous presse.*)
4. Applications à la biologie. Variations continues et Biométrie, par L. BLARINGHEM et C.-E. TRAYNARD.

TOME IV. — *Applications diverses et conclusion.*

1. Applications au tir, par J. HAAG ; 1926................... 35 fr.
2. Applications aux jeux de hasard, par Émile BOREL.
3. Compléments divers.
4. Conclusion : la portée philosophique de la théorie des probabilités, par Émile BOREL.

80368 27. Paris. — Imprimerie GAUTHIER-VILLARS et Cie, 55, quai des Grands-Augustins.

A